THE UNIVERSITY of EDINBURGH

Title	Cosmological parameter estimation from galaxy surveys
Author	Burkey, Daniel.
Qualification	PhD
Year	2004

Digitisation Notes:

- pages iii, iv, vi, 120, 124, 132 missing from original

Cosmological Parameter Estimation from Galaxy Surveys

DANIEL BURKEY

Institute for Astronomy

School of Physics

Doctor of Philosophy

The University of Edinburgh

April 2004

Declaration

I hereby declare that this thesis is not substantially the same as any that I have submitted for a degree or diploma or other qualification at any other University. I further state that no part of my thesis has already been or is being concurrently submitted for any such degree, diploma or other qualification.

This thesis is the outcome of my own work except where specifically indicated in the text.

Daniel Burkey
Edinburgh,
April 2004.

Acknowledgements

I'd like to thank my supervisor Andy Taylor for all his time over the last few years. I still have piles of paper with Andy's explanations of things, which don't involve any words or equations but are mostly squiggly lines and ellipses and other strange shapes. Some I recognise as power spectra, some are error ellipses and some need to be de-projected into a four dimesional parameter space (or something!). I'd also like to thank Will Percival for his help with the spherical harmonics code over the last two years. Most of the work in Chapter 4 involved adapting Will's code for the combined analysis, and he also contributed many ideas for diagnosing problems.

I must thank Tara for inventing "Thesisplot" and then beating everyone to hand in first. Inspirational if slightly annoying! Thesisplot did prove to have a Millenium-bug style flaw in it though – when most of us ran over by more than three months. Also I thank everyone I've shared offices with over the years. In particular I must thank Thomas for always being there. By which of course I mean literally *always being there* – in the office, often asleep on the desk.

My flatmates Ian and Olivia have been tremendous help. Ian paid all of our bills for well over two years (!) – not sure we've settled that by the way – and Olivia is possibly the ideal person for watching telly with. Except football. Comments like " I like it when the little men kiss each other" show there's still room for improvement there.

Finally I'd like to thank my family (and their dogs). Simon has helped by sending lots of people to visit me during the festival! My Grandad has contributed his positivity and optimism, and my Grandma would have been very enthusiastic about my thesis, as she was about everything else. And finally my Mum and Dad have always encouraged me to do a PhD and I hope they realise how important they've been.

Contents

List of Figures

Abstract

Galaxy surveys contain a wealth of information on the distribution of galaxies and dark matter in the Universe. Modern surveys yield large numbers of redshifts giving an unprecedented insight into the clustering of galaxies. The peculiar velocities of galaxies – motions caused by the local gravitational potential – allow the dark matter distribution to be inferred. This thesis deals with the statistical analysis of galaxy redshift and peculiar velocity surveys, concentrating on their exploitation for cosmological parameter estimation. The work is divided into two themes. Firstly predictions are made of the information content of galaxy surveys and the problem of optimising surveys' designs to maximize this information is discussed. The second part is the development of a maximum likelihood method of constraining key parameters which deals most accurately with the spherical nature of modern wide–field surveys. Throughout the thesis, the 6 degree Field Galaxy Survey (6dFGS) is used to demonstrate the methods. Through application of the information theory to the 6dFGS, key predictions are made as to the strengths of the survey. The likelihood method is also developed with the intention that it be used in a future analysis of the 6dFGS.

In the information analysis section, the Fisher information matrix is introduced. It is used to obtain new analytical expressions for the cosmic variance of key cosmological parameters constrained from the galaxy power spectrum of any survey given its volume. A technique for survey optimisation is introduced, which defines a set of parameters to describe the geometry and selection criteria of a galaxy survey. Using reasonable assumptions about the dependency of a survey's timescale on these parameters, it is shown that the survey design may be optimized in the sense of maximizing its information content with respect to a cosmological parameter. The optimal design is calculated for the 6dFGS and predictions are made from the Fisher matrix, of the attainable uncertainties on future cosmological parameter predictions from the survey. Similar analysis is performed for the peculiar velocity survey.

To fully exploit the information content of redshifts and velocities it is necessary to perform a joint analysis of the two. This is particularly suitable for the 6dFGS since it incorporates a redshift and velocity survey which use similar selection critetia. The Fisher matrix for the two data sets of redshifts and velocities is derived. It is then used to predict the information content of the combined surveys, as well as the correlations between the parameters. It is found that the great advantage of combining redshifts and peculiar velocities is that the combination breaks the degeneracy between the redshift distortion parameter and the mass–galaxy correlation coefficient – allowing the simultaneous con-

straint of both.

The second part of the work develops a method of constraining these two parameters. Redshift space distortions are best dealt with by expanding the galaxy distribution of a survey in spherical Bessel functions and spherical harmonics. This separation of angular and radial distributions is also convenient analytically as a way of separating out the window function from the radial selection function. Models are constructed of the combined covariance matrix of the redshift–velocity data set. The models accurately describe the effects of redshift space distortions and stochastic galaxy biassing, to linear order. Software developed to perform this analysis is then tested on a suite of simulations of the 6dFGS.

Finally attention is paid to the problem of optimal data compression. In a real survey with a complicated survey mask, the final covariance matrix can be unfeasibly large to invert at different points in parameter space. The best way of performing this compression is shown to be Generalised Optimal Mode Analysis (GOMA) which optimally compresses the data set in a way that retains the most information on a specific parameter. The results of the maximum likelihood method applied to a compressed data set from the 6dFGS simulations are shown. The results are a good reflection of what will be possible when the survey reaches completion in mid–2005.

1 Introduction: Fundamentals of Cosmology

1.1 Preamble

In the 20^{th} century the Big Bang model became widely accepted as the correct description of the growth and evolution of the Universe. The evidence in its favour is overwhelming. The model is based on the Copernican Principle, the observation that we are not at any special place in the Universe. This together with the observed isotropy of the galaxy distribution implies homogeneity on large scales – the Cosmological Principle. The first major observational result was the discovery by Slipher in the 1910s of the redshift of external galaxies. Then Hubble (1929) found that galaxies receded with velocities proportional to their distance. Combined with the Copernican Principle this implies that observers at any point in the Universe will witness other galaxies receding from them: with themselves the centre of the expansion.

The expanding Universe model was further vindicated by the discovery by Penzias and Wilson (1965) of the Cosmic Microwave Background (CMB). This is a near perfect black–body spectrum coming from the background sky, with a present day temperature of about 2.7K. In the Big Bang model, the early Universe is a hot plasma in which electrons and nuclei are unbound, and photons are continuously Thomson scattering off the electrons. This changes at the time of decoupling when the Universe cools to a temperature where atoms can form. Thomson scattering of the photons ceases and the radiation is free to travel unhindered. The radiation observed in the present day comes from a spherical surface surrounding the observer. This is the surface of last scattering whose radius is the distance travelled by a photon from the time of decoupling to the present. The observed black body temperature is much smaller that that of the plasma at decoupling, since the spectrum is observed redshifted by the expansion. The existence and isotropy of the CMB is strong evidence in favour of a Big Bang – it implies that the Universe was denser in the past.

The Big Bang also helps to solve one of the problems of 19^{th} century physics – Olbers' paradox. In short Olbers' paradox states that if the Universe is spatially infinite and has infinite age, then the night sky should be bright. The lack of a bright night sky implies that there was an epoch of galaxy/star formation, consistent with the Big Bang model.

The predicted age of the Universe found using the Big Bang model and using accepted values for key cosmological parameters, fits in with other limiting estimates. A minimum age for the Earth can be deduced from the fossil record and the ages of older stars in the galaxy have been determined with reasonable accuracy. All sources agree with an order of magnitude value of 10^{10} years.

Probably the most striking success of the Big Bang model is the theory of nucleosynthesis. This accurately predicts the ratios of the observed abundances of light elements. Given the successes of nucleosynthesis, the explanation of the CMB and the observed recession of other galaxies, the Big Bang has become the only feasible model. There are a few problems due to the smoothness of the CMB and the observed flatness of the Universe, for which solutions have been propsed in the form of cosmic Inflation. This will be discussed later on.

1.2 Time, Distance and Horizons

In order to separate out the expansion of the Universe from so–called peculiar motions, it is common to write distances as the product of a scale factor $R(t)$ increasing with the expansion, and comoving units x so that $r(t) = R(t)x$. Hubble's uniform expansion may then be written $v = Hr$ where $H = \dot{R}/R$ is the Hubble parameter. The cosmological time coordinate may be defined as the time at which the mean density of the Universe reaches a certain value. The large scale homogeneity of the galaxy distribution means that such a density will be the same for all comoving observers.

The expanding Universe is best described by the Robertson Walker metric:

$$c^2 d\tau^2 = c^2 dt^2 - R^2(t)[dx^2 + S_k^2(x) d\psi^2] \tag{1}$$

where the function $S_k(x)$ takes the form

$$S_k(x) = \begin{cases} \sin x & (k = 1) \\ \sinh x & (k = -1) \\ x & (k = 0) \end{cases} \tag{2}$$

depending on whether the Universe has positive, negative or no curvature respectively. The Robertson Walker metric may be used to describe the propagation of light which must follow

null geodesics. The radial motion of a photon is then described by

$$dr = cdt/R(t). \tag{3}$$

The integral of dr from $t = 0$ to the present gives the particle horizon – the greatest distance from which light signals can have reached an observer by the present day. The particle horizon depends on the nature of the scale factor – particularly at early times when the Universe may have been exponentially smaller. The expansion also causes emitted photons to be observed with larger wavelengths. The redshift z, defined in terms of emitted and observed frequency can be written

$$\frac{\nu_{emit}}{\nu_{obs}} \equiv 1 + z = \frac{R(t_{obs})}{R(t_{emit})} \tag{4}$$

or in terms of a dimensionless scale factor normalised to the current epoch, $a(t) = (1 + z)^{-1}$.

1.3 Cosmological Parameters

1.3.1 The Friedmann equation.

The Friedmann equation is the equation of motion for the scale factor. The form of the equation can be found from a Newtonian argument in which the motion of a test particle at the edge of a sphere is affected only by the gravity of the mass inside the sphere. Conservation of energy then implies that $(\dot{R}x)^2/2 - GM/(Rx) = \text{constant}$ where $M = \frac{4}{3}\pi(Rx)^3\rho$ is the interior mass, and Rx is the scale factor times some arbitrary comoving radius. Choosing $x = 1$ this becomes:

$$\dot{R}^2 - \frac{8\pi G}{3}\rho R^2 = -kc^2 \tag{5}$$

which is Friedmann's equation. General Relativity is needed to equate the constant k to the curvature of spacetime. Equation (5) implies that the Universe will be flat ($k = 0$) if the density has a critical value $\rho_c = \frac{3H^2}{8\pi G}$. The density parameter:

$$\Omega \equiv \frac{\rho}{\rho_c} = \frac{8\pi G\rho}{3H^2} \tag{6}$$

can then be used to determine the curvature of the Universe. Ω includes all contributions to the density including baryons, dark matter and dark energy, and unless otherwise stated refers to its

value at the current epoch. Another useful definition is the dimensionless present day Hubble parameter, defined as $h \equiv H_0/(100\mathrm{kms}^{-1}\mathrm{Mpc}^{-1})$.

Solutions to the Friedmann equation can be found by defining the equation of state. It is common to assume the present density to be dominated by pressureless matter (dust). Particle continuity implies that $\rho \propto R^{-3}$. If the Universe is also flat ($k = 0$), Friedmann's equation is solved by $R \propto t^{2/3}$ giving the age of the Universe as $t_0 = 2/3H_0$. This is the Einstein – de Sitter model. The more general case of $k \neq 0$ can be solved with parametric solutions for R and t:

$$R = kR_*[1 - C_k(\eta)]$$
$$ct = kR_*[\eta - S_k(\eta)] \tag{7}$$

where R_* is a function of the present day density and Hubble parameters, and C_k and S_k are either cos and sin or cosh and sinh for $k = +1$ and $k = -1$ respectively. The $k = +1$ scenario leads to the recollapse of the Universe under its own gravitational attraction – the so called Big Crunch. The recent detection of a significant vacuum energy spoils this simplified matter–only model. The presence of a cosmological constant term in the Friedmann equations, as discussed below, means that the fate of the Universe is no longer solely dependent upon its matter content.

The mass density of photons, corresponding to their energy, decreases with the expansion of the Universe due to both the increase in proper volume elements and the redshifting of the photons' frequencies. The density of radiation thus scales as $\rho \propto R^{-4}$. With this steeper decline it is inevitable that at some point in the past radiation was dominant over matter. In the early Universe where R is very small the radiation term dominates over the matter and curvature terms: a flat, radiation–dominated Universe is a good approximation. Solving the Friedmann equation for the radiation equation of state gives the scale factor as $R(t) \propto t^{1/2}$.

1.3.2 The Acceleration Equation

For an adiabatic expansion $dE = -pdV$, hence $d[\rho c^2 R^3] = -pd[R^3]$. Using this and differentiating the Friedmann equation gives the acceleration of the scale factor in terms of density and pressure:

$$\ddot{R} = -4\pi GR(\rho + 3p/c^2)/3. \tag{8}$$

It is clear from this equation that the Universe is decelerating unless its dominant component has a negative pressure. In fact recent observations of supernovae have shown that the Universe is accelerating in its expansion. This indicates that some sort of constant density component – a vacuum energy – is required. The pressure of a constant density undergoing adiabatic expansion is $p_{vac} = -\rho_{vac}c^2$ so that the term in brackets in equation (8) is negative and the Universe is accelerating. If the vacuum term dominates the energy budget then solving the Friedmann equation for constant density gives an exponentially increasing scale factor for $k = 0$. This is exactly the form needed in the Inflation scenario in the early Universe (see below). The vacuum term contributes a $+\Lambda/3$ term at the end of equation (5).

1.3.3 Parameterizing the relative contributions

Density parameters may be defined for the individual components: $\Omega_m = \rho_m/\rho_c$, $\Omega_r = \rho_r/\rho_c$ and $\Omega_v = \rho_v/\rho_c$ at the present day. A physically interesting epoch is the time of matter radiation equality. In terms of present day values for matter and radiation density this is given by $(1 + z_{eq}) = \rho_m/\rho_r$. Using Stefan's law and the definition of the matter density parameter this gives:

$$1 + z_{eq} \propto \Omega_m h^2 T^{-4} \tag{9}$$

where the radiation temperature is that of photons in the present day. The photon contribution can be found very accurately from the CMB. We are then left with the parameter $\Omega_m h^2$ (or often the parameter $\Omega_m h$ when using velocity units) to describe the epoch of matter radiation equality.

Given the different dependencies of these components on the scale factor it is possible to write a general equation of state at any epoch as:

$$\rho = \rho_c \left(\Omega_v + \frac{\Omega_m}{a^3} + \frac{\Omega_r}{a^4} \right). \tag{10}$$

This equation may be used to obtain the look back time to an object of given redshift or its proper distance. From the Friedmann equation, the Hubble parameter at redshift z is:

$$H(z) = H_0[\Omega_r(1+z)^4 + \Omega_m(1+z)^3 + \Omega_v + (1-\Omega)(1+z)^2]^{1/2} \tag{11}$$

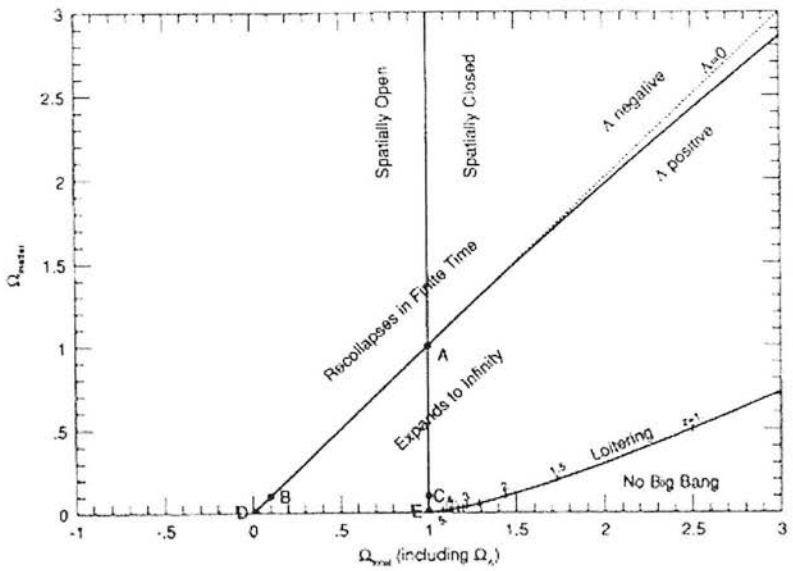

Figure 1: *Fate and curvature of the Universe as a function of the density contributions (taken from Carroll et al. (1992)).*

where Ω is the total density parameter. The lookback time to a redshift z is then:

$$t(z) = \int_0^z \frac{1}{(1+z)H_0}[\Omega_r(1+z)^4 + \Omega_m(1+z)^3 + \Omega_v + (1-\Omega)(1+z)^2]^{-1/2}dz. \quad (12)$$

Similarly the proper distance is

$$r(z) = \frac{c}{H_0}\int_0^z [\Omega_r(1+z)^4 + \Omega_m(1+z)^3 + \Omega_v + (1-\Omega)(1+z)^2]^{-1/2}dz \quad (13)$$

which is the distance measured by a chain of rulers from the observer to the object. By setting the redshift to infinity in these equations we can find the age of the Universe and the particle horizon respectively.

The set of cosmological parameters H_0, Ω_r, Ω_v and Ω_m collectively give the age of the Universe, its geometry, its eventual fate and its composition, as well as providing the tools for obtaining distances and lookback times to galaxies given their redshift. Figure 1 shows the different possibilites. The current concordance model favours $\Omega_m \approx 0.3$, $\Omega_v \approx 0.7$, $H_0 = 72 \mathrm{kms}^{-1}\mathrm{Mpc}$ and $\Omega_r h^2 \approx 10^{-5}$.

1.3.4 The Cosmological constant problem

Particle physics allows the cosmological constant to be estimated. The concordance model predicts a flat Universe in which:

$$1 = \Omega_m + \Omega_\Lambda + \Omega_k. \tag{14}$$

Particle–antiparticle pairs are constantly being created and destroyed by virtue of the Heisenberg uncertainty principle. Thus as Ω_m changes in equation (14) so does Ω_Λ. Ω_m quantifies the density of real particles and Ω_Λ quantifies that of virtual particles. The vacuum has an energy density ρ_{vac} (Zeldovich (1967)) due to the gravitational influence of the virtual particles. Particle physics theory allows one to estimate the value of ρ_{vac} but unfortunately the estimates disagree with observational limits by a factor of 10^{120}. This cosmological constant problem is one of the most important problems in modern physics.

1.4 The Distance Ladder – measuring H_0

Accurate distances from redshift independent methods are required to calculate the Hubble parameter. At the smallest end of the distance scale, the parallax of nearby star clusters can be used to obtain distances- this has been done most recently by the HIPPARCOS satellite (van Leeuwen (1996)). Distances to galaxies may be obtained through the observation of Cepheid variable stars. The period of these stars is known to correlate with their luminosity; a prediction of luminosity combined with apparent magnitude gives the distance to the star. The most common use of Cepheid variables has involved the nearby Large Magellanic Cloud. The Hubble Space Telescope Key Project (Freedman et al. (1997)) has gone further, using Cepheids in a number of nearby galaxies, covering distances out to $\sim$ 20Mpc. To obtain distances on larger scales requires the use of other distance indicators. Distance indicators for use in large galaxy surveys are discussed in depth in Chapter 2 and so their discussion is deferred until then. Attempts to constrain H_0 however, have concentrated on the use of type Ia supernovae as standard candles. A correlation has been found between the time taken to reach maximum light and the luminosity of the supernova. The accuracy of supernovae as distance indicators is unsurpassed allowing distances to be measured to $\sim 6\%$ precision. The method involves the use of Cepheid distances at low redshift for calibration and values of $H_0 = 72 \pm 8 \mathrm{kms}^{-1}\mathrm{Mpc}$

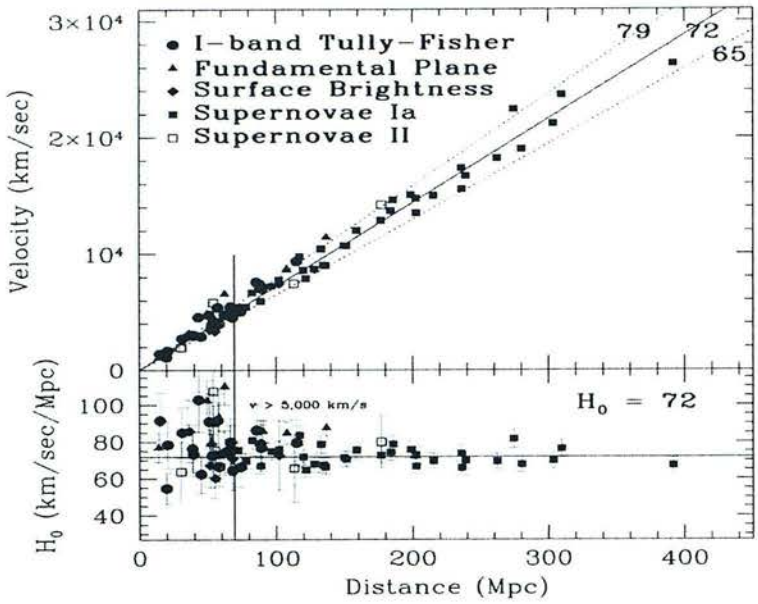

Figure 2: *The Hubble diagram found from the HST Key Project (taken from Freedman et al. (2001)). Different distance indicators are used: they are all calibrated accurately to Cepheid distances at low redshift. The most accurate estimates come from Type Ia supernovae.*

have been obtained (Freedman et al. (2001)). Figure (2) shows results from the HST Key Project using the Cepheid distances to calibrate a number of distance indicators.

1.5 Observational constraints on Ω_Λ

The most compelling evidence for a non–zero cosmological constant has been the use of Type Ia supernovae to obtain redshift independent distances. The luminosity distance to an object d_L – i.e. the distance at which the flux follows the Euclidean law $f \propto L/4\pi d_L^2$ – is dependent on the values of cosmological parameters in a similar way to the proper distance. The distribution of apparent magnitude against redshift, if precise enough, can differentiate between cosmological models and so constrain Ω_Λ. Galaxy surveys can also be used to detect the cosmological constant. Alcock and Paczynski (1979) suggest a geometric test for Ω_Λ: the assumption of an incorrect geometry leads to an effective squashing of space along the line of sight – leading to an anisotropy in clustering statistics. A more general way of parameterising dark energy is to use the equation of state parameter w where $p = w\rho c^2$. The cosmological constant provides

$w = -1$, although more complicated forms of quintessence are not yet ruled out. Lahav (2002) reviews the ability of redshift surveys to measure dark energy through various methods, finding that although promising, most methods suffer from degeneracy with other parameters such as the matter density and galaxy biassing.

1.6 Big Bang Nucleosynthesis

As the Universe was denser in the past so it was hotter. There was a time in which protons and neutrons were free particles but were non–relativistic. The reactions of the particles at this time, are found to have been faster than the expansion rate and so the Universe was in thermal equilibrium. The number densities of neutrons and protons can be found from the Maxwell–Boltzmann distribution:

$$\frac{N_n}{N_p} = \left(\frac{m_n}{m_p}\right)^{3/2} \exp\left[\frac{-(m_n - m_p)c^2}{k_B T}\right]. \tag{15}$$

Neutrons are heavier than protons and so can decay via a weak interaction:

$$n \rightarrow p + e^- + \overline{\nu}_e + 0.8\text{MeV}. \tag{16}$$

So long as there are significant numbers of photons with energies above 0.8MeV the reaction will occur in both directions and thermal equilibrium holds. At $k_B T = 0.8$MeV equilibrium breaks down and the forward reaction becomes more common. At this stage the number densities of neutrons and protons are roughly:

$$\frac{N_n}{N_p} \approx \exp\left(-\frac{1.3\text{MeV}}{0.8\text{MeV}}\right) \approx \frac{1}{5}. \tag{17}$$

This corresponds to the Universe at a temperature of $10^{10} K$, a few seconds after the Big Bang. The neutron population will continue to decrease until nucleosynthesis when they can be bound with protons to form the most stable light element – He^4. Detailed calculation shows that this occurs at about 3 minutes after the Big Bang when the neutron to proton ratio has decreased further to $N_n/N_p \approx 1/7$. Assuming that all of the available neutrons are bound into He^4 the mass fraction of He^4 is

$$Y_4 = \frac{4(N_n/2)}{N_n + N_p} \approx 0.25. \tag{18}$$

This is one of the great successes of nucleosynthesis in that it explains the observed almost–uniform abundance of He^4 in stars as noted by Gamow in the 1940s. Another important prediction of the nucleosynthesis theory is that it places an upper limit on the baryonic contribution to

the matter density parameter. The observed present day abundance of Deuterium can be used to predict the baryon abundance as

$$\Omega_b h^2 = 0.02 \pm 0.001 \tag{19}$$

(Burles (2002)). If cosmology favours a larger matter density parameter then its constituents must be largely non–baryonic.

1.7 Dark Matter

For the past 70 years it has become increasingly clear that the majority of matter in the Universe is not luminous. Different pieces of evidence from different techniques point to a discrepancy between the mass we can infer and the luminous mass we can detect. Firstly, the rotation curves of galaxies are found to flatten off with increasing radius from the galaxy centre. If the orbits of stars near the edge of a galaxy were purely due to the gravitational attraction of a centralized mass, the orbital velocities would be expected to decline such that,

$$v_{star}^2 = \frac{GM_{galaxy}}{r} \Rightarrow v_{star} \propto r^{-1/2}. \tag{20}$$

Instead, Rubin and Ford (1970) found that the orbital velocities remained almost constant above a certain radius, implying that galaxies are embedded in large haloes of dark matter ($\rho \propto r^{-2}$). Other dynamical evidence comes from observations of bulk galaxy motions in clusters. This was first seen by Zwicky (1937) who noticed that velocities of galaxies in the Coma cluster were so large that the cluster would need an enormous mass in order to be gravitationally bound – far more than could be present in stars.

Further evidence comes from observations of gravitational lensing – the bending of light rays in a gravitational field. This phenomena can be seen on a wide variety of scales. Microlensing (i.e. lensing by what may be approximated as a point mass) has been observed by studies of the light curves of stars in the Large Magellanic Cloud (LMC). The MACHO project and others have observed millions of stars in the LMC in order to witness a few lensing events when the line of sight to the star is crossed by some unknown dark matter object in the halo of our own galaxy – a MAssive Compact Halo Object. The fact that this has been witnessed at all is evidence for the existence of dark matter, although some of these events may have been the

result of lensing by stars. On larger scales galaxies themselves can act as the lens, lensing high redshift galaxies and quasars. On a yet larger scale, galaxy *clusters* can act as lenses of very distant galaxies. When the data is collected and maps made of the inferred mass distribution of these lenses, the mass is found to be both larger and distributed differently to the luminous mass. A striking example of this is the lensing galaxy cluster Abell 2218 where the inferred lensing mass is found to be two orders of magnitude greater than the total luminous mass of the cluster members.

1.7.1 Different types of dark matter

Given the upper bound on baryon density provided by nucleosynthesis and the currently favoured value of $\Omega_m \approx 0.3$ from the break scale in the power spectrum, a significant amount of the dark matter must be non–baryonic. A number of different scenarios are possible but they tend to be variations or mixtures of Hot Dark Matter (HDM) and Cold Dark Matter (CDM). In 1966 Zeldovich proposed non-zero mass neutrinos as possible HDM candidates. In this scenario the neutrinos are relativistic at the time of decoupling and have the effect of damping density fluctuations. This invokes a picture of top down structure formation where structures in the density field form initially on the largest scales and then break up into smaller and smaller pieces. Qualitatively, because neutrinos only interact weakly they can 'free stream' out of regions of overdensity into regions of underdensity thereby wiping out small scale perturbations. The size of early structures therefore has a lower limit which is a function of the horizon size at the time when the particles become non relativistic. The existence of early galaxies at redshifts of $z \sim 6$ cause problems for HDM models since free streaming makes it difficult for galaxies to have formed so soon.

An alternative to HDM is cold dark matter. Cold in this context means non–relativistic. This invokes a scenario in which the smallest scale structures form initially, gradually building up into larger and larger perturbations. Attempts to directly detect dark matter particles are being made by observing the recoil of nucleons interacting with the dark matter. This involves careful removal of thermal and environmental effects and experiments are often conducted in deep mines at very low temperatures. The neutralino, which is the lightest particle predicted by supersymmetry, is the most promising CDM candidate in particle physics. Although CDM

is currently the favoured model, many other dark matter candidates exist – some consist of mixtures of HDM and CDM – but they are not favoured by current data.

Baryons can account for a small proportion of the dark matter. Brown dwarfs are stars whose mass is too low for hydrogen burning to occur. It is thought that they contribute to at most 20% of the MACHOS observed in the lensing events discussed earlier. Of course brown dwarfs cannot be the only form of dark matter because of the nucleosynthesis constraint. One possibility which gets around nucleosynthesis is to have the dark matter trapped in primordial black holes. 'Primordial' implies that they were formed before nucleosynthesis and so the baryonic mass contribution could have been greater before this epoch.

1.8 The Cosmic Microwave Background

The CMB is stikingly uniform across the sky – with anisotropies detected as small temperature fluctuations in CMB maps of order 10^{-5} of the mean. These variations in temperature correspond to density variations in the decoupling plasma, and in the gravitational instability paradigm it is these initial density perturbations which grow through gravitational attraction into the galaxies we see today. The anisotropies in the observed temperature are small and the mean black body distribution has a temperature of:

$$T = 2.725 \pm 0.002K \tag{21}$$

(Mather et al. (1999)). The existence of the CMB is one solution to Olbers' paradox: the night sky is uniformly glowing but its spectrum is redshifted beyond the optical range.

1.8.1 The origin of Black Body Radiation

The spectrum of the CMB is that of an almost perfect black body. Black body radiation is a natural consequence of the equipartition of thermal energy. A body in thermal equilibrium absorbs all radiation incident upon it and then re–radiates the energy in a spectrum given by Planck's law:

$$u(\nu, T) = \frac{8\pi h\nu^3}{c^3} \frac{1}{e^{h\nu/kT} - 1} \tag{22}$$

for the energy density $u(\nu, T)$ as a function of frequency, which depends only upon the body's temperature. The theory of nucleosynthesis, and its predictions, as discussed in Section (1.6), rely on the primordial plasma being in a state of thermal equilibrium: the Maxwell–Boltzmann distribution applies in this regime. Thermal equilibrium is itself predicted at early Universe temperatures since the reactions of particles would have been faster than the expansion rate at this time. The blackbody nature of CMB radiation is thus an important piece of evidence towards the Big Bang model. It is difficult to imagine a non–cosmological source for the CMB as would be required for instance by the Steady State theory.

1.8.2 CMB Anisotropies

One anisotropy which must be discarded is the dipole anisotropy due to the Earth's peculiar motion (with respect to a frame in which the Universe is expanding isotropically). The other CMB anisotropies are characterized as primary or secondary. Primary anisotropies were formed at the time of decoupling. There are three important primary anisotropies. Sachs–Wolfe perturbations arise due to gravitational redshifting of photons – an effect which is greater in regions of high gravitational potential. Photons coming from regions of low potential are blueshifted. Another primary anisotropy is caused by intrinsic perturbations. Dense regions are hotter and recombine later than underdense regions. This means that overdense regions are less redshifted than underdense regions since they have recombined closer to the present day. Finally Silk damping suppresses fluctuations on small angular scales. Silk damping is caused by the diffusion of photons through the plasma.

The secondary anisotropies occur in the radiation en route from the surface of last scattering. One such phenomenon is the Rees–Sciama effect. If a photon passes through a gravitational potential en route to the observer, then it will be gravitationally blueshifted on passing into the potential and gravitationally redshifted on its way out. The two effects will counteract leaving no net change. But if the gravitational potential changes during the time of passage of the photon then there will be a net shift in the photon's frequency. Another secondary anisotropy is caused by the Sunyaev Zeldovich effect – the inverse Compton scattering of photons en route from the surface of last scattering. This effectively removes photons from the microwave region of the spectrum. Clearly this depends on the number of electrons encoun-

tered by photons when travelling from different parts of the sky and is therefore an anisotropic process.

1.8.3 Cosmological parameters from the CMB

The spherical nature of the CMB dictates that the temperature fluctuations should be expanded in a spherical basis. The autocorrelation function is defined as

$$C(\theta) = \langle \frac{\delta T}{T}(\alpha) \frac{\delta T}{T}(\alpha + \theta) \rangle. \tag{23}$$

This can be written in terms of Legendre polynomials $P_l(\theta)$ as

$$C(\theta) = \frac{1}{4\pi} \sum_{l=2}^{\infty} \langle |a_l^2| \rangle (2l+1) P_l(\cos\theta). \tag{24}$$

The $l = 0$ multipole is the mean temerature of the sky and the $l = 1$ multipole is the dipole anisotropy. Higher values of l correspond to fluctuations on scales of

$$\theta \approx \frac{180^\circ}{l}. \tag{25}$$

The powers $C_l \equiv \langle |a_l^2| \rangle$ are adjusted to get the best fit to the observed temperature distribution. The distribution of C_l against l is the standard way of displaying the power spectrum of the CMB.

Figure (3) shows the CMB power spectrum observed by the Wilkinson Microwave Anisotropy Probe (WMAP, Spergel et al. (2003)) and the best fit cosmological model. The flat region of the power spectrum on large angular scales was first observed by the Cosmic Background Explorer satellite (COBE, Smoot (1999)) and corresponds to scales above the sound horizon at last scattering where only scale invariant gravitational perturbations play a role. Since the sound horizon can be calculated at the time of last scattering, the angle on which the intrinsic perturbations kick in – the position of the first peak – can be predicted. The prediction requires a model for the curvature of the Universe which affects observed angular sizes. The position of the first peak has been used in this way by WMAP and previous observations to predict a flat cosmology consistent with independent predictions from redshift surveys combined with the supernova project.

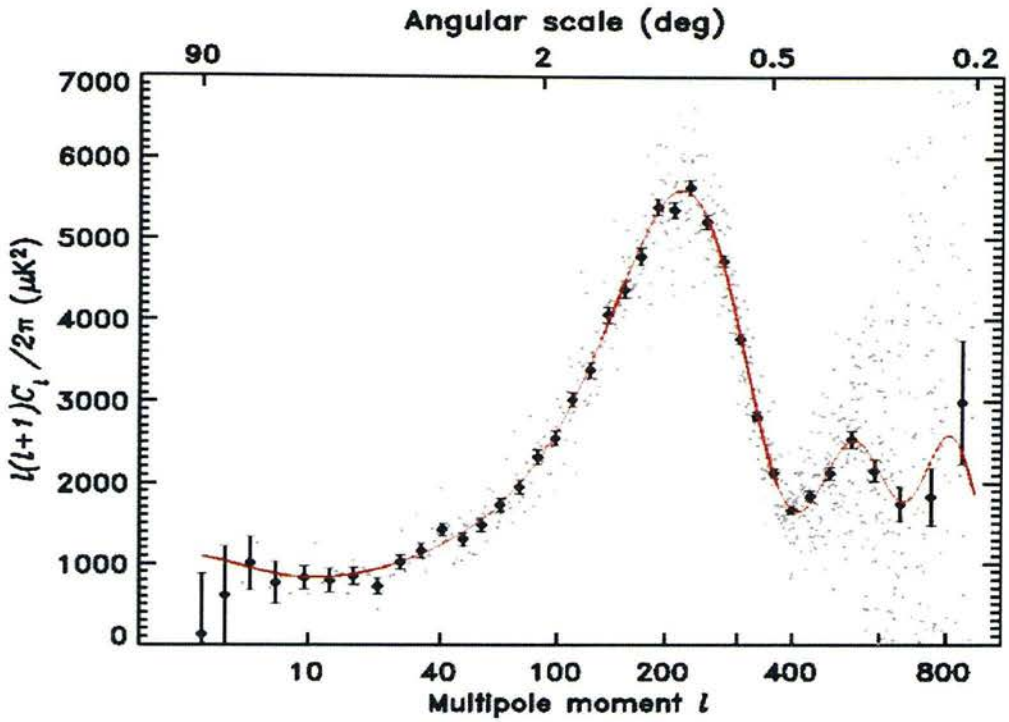

Figure 3: *The power spectrum of the CMB as observed by WMAP (taken from Spergel et al. (2003)). The red line is the best fit ΛCDM model.*

The succesive peaks in Figure (3) are the result of the acoustic oscillations of the primordial plasma 'freeze–framed' at the time of decoupling. The second peak is a rarefaction of the plasma whereas the first and third peaks are compressions. Models of these acoustic oscillations show that the relative height of the 2nd peak with respect to the first or third peaks, is sensitive to the baryon density Ω_b. Baryon loading is the effect of the weight of baryons on the acoustic oscillations. The baryons cause the rarefactions in the oscillating plasma to be less pronounced than the compressions. Spergel et al. (2003) find $\Omega_b = 0.047 \pm 0.006$. The third peak is needed to constrain the amount of Dark Matter. The amplitude of the power spectrum increases with the dark matter content of the Universe, but three peaks are required in order to distinguish this effect from baryon loading. Detailed modelling of the primordial plasma has allowed WMAP in conjunction with redshift surveys and supernovae to constrain a host of different cosmological parameters as described in Spergel et al. (2003).

The damping tail of the CMB power spectrum can be used as a consistency check for these parameters. Power is damped at very short wavelengths by photons streaming from overdensity to underdensity destroying the temperature variations. This occurs at scales below the mean

free path for photons before recombination which in turn depends upon the dark matter content and the baryon fraction. The observed damping scale also depends upon the curvature in the same way as the first peak. The constraints from these effects will provide an independent corroboration of the results from the first three peaks and will be achievable with the upcoming Planck satellite (Lawrence and Lange (1997)).

1.9 Inflation

1.9.1 Problems with the Big Bang

Cosmological Inflation describes a period in the early Universe, before radiation domination, in which the equation of state was very different, allowing a period of rapid expansion to occur. Inflation is motivated by a number of problems with the Big Bang model. The first is the flatness problem. The current best estimates for the present day density parameter put it very close to one. The Friedmann equation can be written:

$$|\Omega - 1| = \frac{|k|}{a^2 H^2} \tag{26}$$

so if $\Omega = 1$ the curvature constant is zero and remains so for all time. Since $a^2 H^2 = \dot{a}^2$ then for both matter dominated and dust dominated scenarios the right hand side of equation (26) is an increasing function of time. For Ω to be close but not equal to unity at the present day, the early Universe must have been extremely flat. Inflation provides a mechanism for this.

Another motivation for Inflation is the horizon problem. The comoving particle horizon can be calculated from:

$$r_H = \int \frac{cdt}{R(t)}. \tag{27}$$

If the early Universe is always radiation–dominated, then $R(t) \propto t^{1/2}$ and $r_H \to 0$ as $t \to 0$. The particle horizon defines the limiting volume in which information about the initial conditions of the expansion could have travelled. This makes it difficult to see how the large scale homogeneity seen at later stages could have been imposed. Detailed calculations show that points on the surface of last scattering which are more than 2 degrees apart cannot have been causally connected by the time of recombination. In this case the plasma off which the CMB photons last scattered could not have reached thermal equilibrium before decoupling.

The solution provided by an inflationary expansion is that a small causally connnected region can become exponentially larger in a small amount of time. To get around the flatness problem we need $R \propto t^{\alpha}$ with $\alpha > 1$ so that in equation (27) the particle horizon does not tend to zero at early times.

Other motivations for Inflation include an explanation of the initial conditions leading to present day structure. Inflation allows quantum fluctuations to be hugely amplified suggesting one possible origin. Anther motivation is the relic problem. Particle physics suggests that many relic particles should have been produced in phase transitions. If the Inflation scenario is correct, the lack of observations of these particles could be because they have been 'diluted' by the expansion.

1.9.2 Solutions provided by Inflation

If the scale factor takes the form $R \propto t^{\alpha}$ then we require $\alpha > 1$ to solve the horizon problem. The acceleration equation (equation (8)) then implies that:

$$\rho c^2 + 3p < 0. \tag{28}$$

The equation of state during inflation must therefore provide a negative pressure. The only component of the three discussed which does this is the vacuum energy. The de Sitter solution:

$$R \propto \exp(\sqrt{8\pi G\rho_{vac}/3})t \tag{29}$$

to the Friedmann equation for a flat vacuum dominated universe, is an inflationary model in which the scale factor tends to a constant as $t \to 0$. In this case there is no Big Bang. The model solves the horizon problem since the horizon is no longer zero at $t = 0$. The flatness problem is also solved since $|\Omega - 1|$ now decreases with time during the inflationary period. This mechanism drives the Universe so close to flatness that all of the expansion post–inflation is not enough to significantly move Ω away from one.

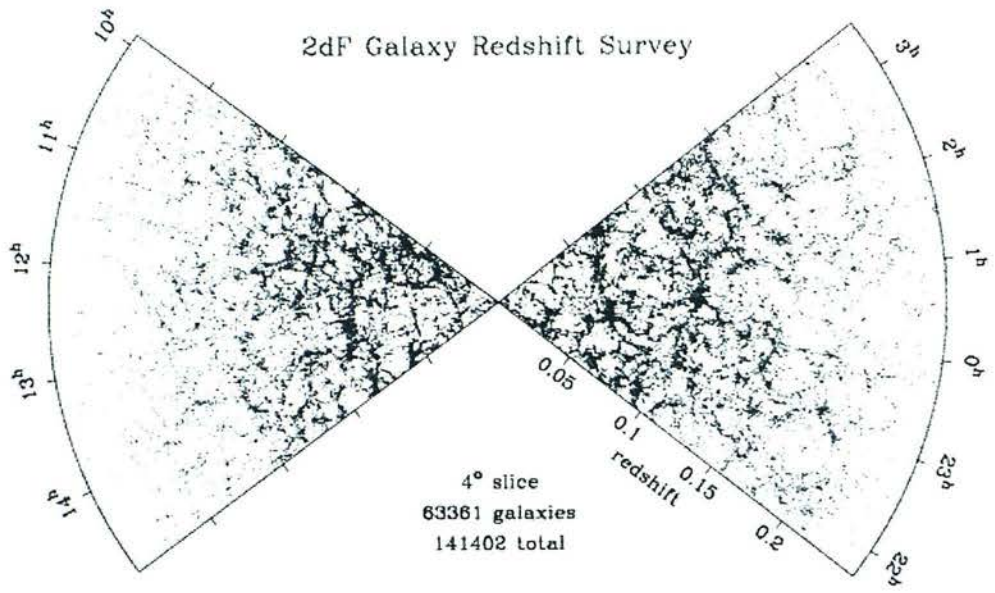

Figure 4: *The galaxy distribution in the 2dF Galaxy Survey.*

1.10 Large Scale Structure

Figure 4 shows the distribution of galaxies observed in the 2dF Galaxy Redshift Survey (Colless et al. (1999)). The clustering pattern of galaxies in the local Universe contains a wealth of information on cosmological parameters, being the result of the gravitational instability of the small perturbations observed in the CMB. The way in which these inhomogeneities have grown since decoupling provides an insight into the expansion of the Universe as well as the nature of galaxy formation.

The remaining part of this introduction sets out the groundwork for the study of large scale structure. Sections 1.10.1 and 1.10.2 show that density perturbations can grow through gravitational instability. Section 1.10.3 derives the growth rate of these inhomogeneities. The Mészáros effect is explained in Section 1.10.4. Sections 1.10.5 and 1.10.6 discuss the effects of pressure gradients and the contribution of baryons. Finally Section 1.10.7 relates the over-density field to the peculiar velocity field.

1.10.1 Linear Perturbation Theory

Models of perturbations in the primordial plasma fall into two classes – abiabatic and isocurvature. In isocurvature models, peaks in the matter density correspond to troughs in the radiation density and vice versa, so $d\rho_m = -d\rho_r$. The CMB power spectrum is not consistent with the isocurvature models and so they will not be discussed further. In adiabatic models the radiation component is perturbed equivalently to the matter component. It is common when describing the density field to use the overdensity δ defined as:

$$\delta(r) \equiv \frac{\rho(r) - \rho_0}{\rho_0}, \tag{30}$$

where ρ_0 is the mean. For adiabatic perturbations we have $\delta_r = \frac{4}{3}\delta_m$.

The growth of structures in the density field is a balancing act between different effects. Perturbations (assumed initially small) grow through gravitational interactions which compete with the effect of the Hubble expansion which tries to pull them apart again. Total collapse is also hindered by pressure forces for the baryons. The limiting wavelength λ_J below which baryonic perturbations will not grow, is called the Jeans length.

In linear perturbation theory each component is treated separately. The components are assumed to obey ideal fluid equations – the continuity equation and a force law describing the effects of gravity and pressure support. Each perturbation is described as a superposition of its Fourier modes. If the wavelength of a mode is outside the Hubble radius at time t, then the evolution of the mode must be analysed with a full relativistic treatment. Similarly during radiation domination relativity is needed to describe the pressure of radiation. During matter domination however, if the mode's wavelength is well within the Hubble radius then a Newtonian treatment suffices as an approximation.

1.10.2 Newtonian Approximation

If the components are ideal fluids they will obey the continuity equation:

$$\dot{\rho} + \rho \nabla . \mathbf{v} = 0 \tag{31}$$

where the dot denotes a convective derivative $\frac{D}{Dt}$ such that

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \mathbf{v} \cdot \nabla. \tag{32}$$

All dotted terms hereafter denote convective derivatives and gradients are in the comoving frame. The comoving acceleration of an element of the fluid is the result of pressure gradients and gravity. Writing the gravitational force on a fluid element as the gradient of a potential term Φ we obtain:

$$\dot{\mathbf{v}} = -\nabla \cdot \Phi - \frac{1}{\rho} \nabla \cdot p. \tag{33}$$

This is Euler's equation. The convective derivative takes into account the contorting of the volume's shape due to these forces. An unperturbed density field may be defined which is perfectly homogeneous. Homogeneity in the density also implies the absence of pressure gradients. This describes the very early stages of the Universe. For this unperturbed field the continuity equation gives:

$$\dot{\rho_b} = -3H_b\rho_b \tag{34}$$

where subscript b denotes unperturbed values of the density and Hubble parameter. A perturbation $\rho = \rho_b + \delta\rho$ in the density field will create a gradient in the gravitational potential and this gravitational attraction will have a small effect on the Hubble expansion. The effect can be described as a small perturbation $H = H_b + \delta H$ in the Hubble parameter. Substituting these changes into equation (31) and discarding non–linear terms in δH and $\delta\rho$ gives:

$$\dot{\rho_b} + \dot{\delta\rho} \approx -3H_b\rho_b - 3H_b\delta\rho - 3\rho_b\delta H. \tag{35}$$

Subtraction of the unperturbed equation gives:

$$\dot{\delta\rho} = -3H_b\delta\rho - 3\rho_b\delta H. \tag{36}$$

Writing this in terms of the overdensity $\delta(r)$ gives:

$$\dot{\delta} = -3\delta H. \tag{37}$$

There are three coordinate systems in common use in Cosmology. Eulerian coordinates are measured with respect to a reference frame fixed in inertial space. Comoving units are measured with respect to a frame expanding with the scale factor and are convenient to describe effects free from the Hubble expansion. Lagrangian coordinates are such that the coordinates

of a set of test particles do not change. In this case the frame contorts to allow for both the Hubble expansion and the effects of gravity and pressure. The Lagrangian system is useful for describing second order effects. The analysis here will concentrate on Eulerian perturbation theory.

The Eulerian position of a test particle in the expanding Universe is just $\mathbf{r} = a(t)\mathbf{x}$ where $\mathbf{x}$ is the comoving position. The proper velocity is then $\dot{\mathbf{r}} = H\mathbf{r} + a\dot{\mathbf{x}}$ which is the combination of the Hubble expansion velocity and a peculiar velocity. Differentiating again gives:

$$\ddot{\mathbf{r}} = a\dot{\mathbf{u}} + 2\dot{a}\mathbf{u} + \frac{\ddot{a}}{a}\mathbf{r} = \mathbf{g}_0 + \mathbf{g} \tag{38}$$

where the acceleration has been split into the unperturbed acceleration of an expanding homogeneous density field $\mathbf{g}_0$, and an acceleration due to the perturbing influence of density fluctuations $\mathbf{g}$; known as peculiar gravity. The peculiar velocity is $\mathbf{u}$. At this stage pressure gradients are neglected. Subtracting the unperturbed terms from equation (38) leaves:

$$\dot{\mathbf{u}} + 2\frac{\dot{a}}{a}\mathbf{u} = \frac{\mathbf{g}}{a}. \tag{39}$$

The second term is the Hubble drag and is the result of observing motions in an expanding Universe. It represents the fact that as the Universe expands the recessional velocity of other galaxies increases. The final differential equation for δ can be found by taking the divergence of equation (39) and using equation (37) giving:

$$\ddot{\delta} + 2\frac{\dot{a}}{a}\dot{\delta} = -\frac{1}{a}\nabla^2\Phi \tag{40}$$

which recognises the fact that $3\delta H = \nabla \cdot \mathbf{u}$. The gravitational potential perturbation is related to the density through Poisson's equation: $\nabla^2\Phi = 4\pi G a \rho_0 \delta$. The solutions to equation (40) in the limit of a flat matter dominated Universe, neglecting pressure, are $\delta(t) \propto t^{\,2/3}$ and $\delta(t) \propto t^{-1}$. The mode which decreases with time will not be important in the present day. The growing mode increases in the same way as the scale factor in this instance. The growing solution describes how small perturbations grow in time through gravitational instability. This is an adequate description of pressureless dark matter but not radiation for which pressure gradients will have an important effect. Since baryons are coupled to radiation through Thomson scattering, baryons are also affected by pressure terms.

1.10.3 The growth rate of perturbations

Equation (37) can be rewritten:

$$\nabla \cdot \mathbf{u} = -\dot{\delta}, \tag{41}$$

the solution to which is

$$\nabla \cdot \mathbf{u} = -f(\Omega_m, \Omega_\Lambda) H \delta \tag{42}$$

where the function

$$f(\Omega_m, \Omega_\Lambda) \equiv \frac{d \log \delta}{d \log a} \tag{43}$$

is the growth rate of perturbations. Lahav et al. (1991) discuss various limiting solutions for the growth rate. The general method to find $f(\Omega_m, \Omega_\Lambda)$ involves solving equation (40) for δ as a function of the scale factor a and then using equation (43). Pressure gradients can be neglected since for most of the time in which the perturbations grow, the Universe is dominated by pressureless dust. The general solution for δ was shown by Heath (1977) to be:

$$\delta \propto H_0^{-2} X^{1/2} a^{-1} \int_0^a X^{-3/2} da, \tag{44}$$

where $X \equiv 1 + \Omega_m(a^{-1} - 1) + \Omega_\Lambda(a^2 - 1)$. For reasonable values of Ω_Λ and Ω_m, Lahav et al. (1991) find the approximate solution:

$$f(z = 0) \approx \Omega_m^{0.6} + \frac{1}{70}\Omega_\Lambda \left(1 + \frac{1}{2}\Omega_m\right). \tag{45}$$

Given the very weak dependence on the cosmological constant, surveys of the local Universe at low redshift have not yet provided constraints on Ω_Λ. For analysis at low z it has therefore been common to adopt the approximation by Peebles (1980) of

$$f(\Omega_m, \Omega_\Lambda, z \approx 0) \approx \Omega_m^{0.6}. \tag{46}$$

Equation (45) will be of fundamental importance in the following chapters as it provides a means by which the matter density parameter and vaccum energy may be constrained in redshift and peculiar velocity surveys.

1.10.4 The Mészáros effect

To study the growth of matter perturbations during radiation domination it is useful to approximate the radiation density as a homogeneous background. The Friedmann equation gives:

$$H^2 = 8\pi G(\rho_m + \rho_r)/3. \tag{47}$$

Changing variables to $y \equiv \rho_m/\rho_r = a/a_{eq}$ allows the perturbation equation to be rewritten as:

$$\delta'' + \frac{2+3y}{2y(1+y)}\delta' - \frac{3}{2y(1+y)}\delta = 0 \tag{48}$$

where the primes denote derivatives with respect to y. This has a growing solution:

$$\delta \propto y + 2/3. \tag{49}$$

This result shows that at the extremes of radiation domination, when $\rho_r \gg \rho_m$, density perturbations are constant, and only when the matter density becomes comparable to the radiation density around matter radiation equality do mass perturbations grow again. This is the Mészáros effect. Perturbations can grow above the causal horizon. When they enter the horizon during radiation domination their growth ceases only to begin again at matter radiation equality. The scale of a perturbation dictates how much time it spends within the particle horizon during radiation domination, and therefore how much its growth is supressed. As discussed in the next section, this leads to a characteristic break scale in the power spectrum of mass fluctuations, which may be used to constrain cosmological parameters.

1.10.5 The Jeans Scale

A useful means of analysing the density field is by decomposing it into its Fourier modes:

$$\delta(\mathbf{x}, t) = \int \frac{d^3\mathbf{k}}{(2\pi)^3}\delta_{\mathbf{k}}(t)\exp(i\mathbf{k}.\mathbf{x}). \tag{50}$$

This has the benefit that physical effects can be clearly separated according to scale, and that the operator ∇ becomes the multiplicative factor $i\mathbf{k}$.

For fluids with significant pressure, the density perturbations will lead to pressure perturbations dp and an extra term will be needed on the right hand side of equation (40):

$$\ddot{\delta} + 2\frac{\dot{a}}{a}\dot{\delta} = 4\pi G\rho_0\delta + \frac{\nabla^2 dp}{a^2\rho_0}. \tag{51}$$

Substitution of equation (50) gives the equation in Fourier space as:

$$\ddot{\delta} + 2\frac{\dot{a}}{a}\dot{\delta} = \delta(4\pi G\rho_0 - c_s^2 k^2/a^2), \tag{52}$$

where the sound speed is defined as

$$c_s^2 \equiv \frac{\partial p}{\partial \rho}. \tag{53}$$

There is therefore a limiting wavelength below which the right hand side of equation (52) is negative and overdensity modes will not grow. This is the Jeans Length:

$$\lambda_J \equiv c_s\sqrt{\frac{\pi}{G\rho}} \tag{54}$$

and it corresponds to the scale on which pressure gradients counterbalance gravity and halt the growth of perturbations.

1.10.6 The effects of baryons

Baryons differ from dark matter particles in that before recombination baryons are coupled to the radiation field through Thomson scattering. This gives the baryons a pressure in contrast to the pressureless dark matter. This means that in the period between matter–radiation equality and decoupling, baryonic perturbations on scales below the Jean's scale, will be damped by pressure waves. On larger scales pressure can be neglected and baryonic perturbations simply follow the same equation of growth as the dark matter. On smaller scales the acoustic oscillations of the baryons cause an overall suppression of power and introduce small oscillations into the power spectrum. Another interesting effect is that of velocity overshoot. A simple model envisages the baryonic perturbations as oscillating as a sound wave in the primordial plasma until recombination, at which point the pressure is instantly turned off. In this model the continuing build up of baryonic perturbations after recombination depends on the phase of the sound wave at the instant the pressure is turned off. Those wavelengths where the overdensity is zero at recombination, consist of sound waves in a pure velocity state. These velocities cause the build up of structure post-recombination. Conversely the modes whose phases were at maximum overdensity are wiped out after decoupling.

1.10.7 The peculiar velocity field

A similar Fourier decomposition of the peculiar velocity field is possible. In Fourier space, equation (42) can be written:

$$\mathbf{u} = -ifH\delta_k \frac{\mathbf{k}}{k^2} \tag{55}$$

showing that the peculiar velocity modes depend linearly on the density modes. This equation relies on the assumption that the velocity field can be written as the gradient of a potential, i.e. that it is curl free. This assumption is valid because vorticity decays due to the expansion. In real space equation (42) can be integrated to give:

$$\mathbf{u}(\mathbf{r}) = \frac{fH}{4\pi} \int d^3\mathbf{r}' \frac{\delta(\mathbf{r}')(\mathbf{r}' - \mathbf{r})}{|\mathbf{r}' - \mathbf{r}|^3}. \tag{56}$$

Assuming a homogeneous solution for the peculiar velocity field $\mathbf{u} = F(t)\mathbf{g}$ gives the peculiar velocity in terms of the peculiar gravity as:

$$\mathbf{u} = \frac{2f}{3H\Omega}\mathbf{g} \tag{57}$$

from the continuity equation and Poisson's equation.

1.11 Statistical Analysis of Observed Fields

1.11.1 The Power Spectrum

The power spectrum is a convenient way to describe clustering strength on different scales in Fourier space. The power spectrum is defined as

$$\langle \delta(\mathbf{k})\delta^*(\mathbf{k}') \rangle = P(|\mathbf{k}|)(2\pi)^3 \delta_D(\mathbf{k} - \mathbf{k}') \tag{58}$$

where the expectation value is taken over an ensemble of realisations of the density field. The ergodic theorem states that this ensemble averaging is equivalent to volume averaging over large causally disconnected regions of space. In equation (58) there is an implicit assumption of statistical isotropy, in that the covariance of the density field is described in terms of a power spectrum dependent upon the magnitude of $\mathbf{k}$ and not its direction. It also assumes statistical homogeneity. Both are required in the Big Bang model. The result is that different modes may be assumed independent in models of structure growth.

1.11.2 The Correlation function

The real space counterpart to the power spectrum is the correlation function, defined as the excess probability over random of finding a neighbouring galaxy at position $\mathbf{r}$:

$$\xi(r) \equiv \langle \delta(\mathbf{x})\delta(\mathbf{x}+\mathbf{r}) \rangle \tag{59}$$

where the average is over positions $\mathbf{x}$. In terms of the power spectrum the correlation function can be written:

$$\xi(r) = \frac{1}{(2\pi)^3} \int d^3\mathbf{k} P(k) e^{i\mathbf{k}\cdot\mathbf{r}}. \tag{60}$$

The central limit theorem implies that the sum of many random variables will exhibit a Gaussian distribution. For this reason it is common to assume that δ will have Gaussian statistics. It is also a prediction of Inflation and a testable characteristic of galaxy surveys that the density distribution is a Gaussian random field. In Fourier space, each individual mode of wavenumber k has a random phase. If $\boldsymbol{\delta}$ is a vector whose components are $\delta(k_i)$ evaluated at a discrete set of N wavenumbers then the probability density function for a Gaussian random field is:

$$p(\boldsymbol{\delta}) = \frac{1}{(2\pi)^{N/2}\sqrt{\det \mathbf{C}}} \exp -\frac{1}{2}\boldsymbol{\delta}\mathbf{C}^{-1}\boldsymbol{\delta}^t. \tag{61}$$

where $\mathbf{C}$ is the covariance matrix of the modes.

The random field hypothesis is testable, since for a purely Gaussian field the statistical information is contained entirely within the first two moments of the distribution. The third moment of a distribution is its bivariance. If the bivariance is non-zero then the Gaussian assumption does not hold. In practice, random phases only hold on scales large enough that non-linear gravitational effects may be neglected. In the power spectrum this corresponds to an upper limit on the wavenumber, for linear analysis. Since the mean of the overdensity is zero by definition, the power spectrum contains all of the statistical information on the field in the linear regime.

1.11.3 Galaxy Biassing

Since the majority of matter is in the form of dark matter, surveys of galaxies' positions do not yield the underlying distribution. It is therefore necessary to make assumptions about the relationship between dark and luminous mass. Some biassing between the distributions may be expected since the galaxy distribution is dependent upon the detailed physics of galaxy formation – local effects such as the composition of the Inter Galactic Medium or the presence of Active Galactic Nuclei are expected to play a part.

The clustering properties of galaxies, separated according to various criteria, has been the subject of much analysis. It is well known that elliptical galaxies are more strongly clustered than spirals. Dressler (1980) studied the galaxy populations in 55 rich clusters and in particular looked at the clustering properties of different morphological types. A well defined relation was found between local galaxy density and galaxy type, indicating an increase in the elliptical and S0 populations with increasing density and a corresponding decrease in spiral populations. More recently, Zehavi et al. (2002) have analysed clustering properties of 29,300 galaxies in the Sloan Digital Sky Survey. They calculated the real space correlation function for the whole sample and then separated the galaxies according to colour – finding that red galaxies exhibit a steeper correlation function than blue galaxies. A similar difference was found when separating the galaxies according to high/low profile concentration or high/low surface brightness. They also obtained a clear measurement of scale independent luminosity bias. Norberg et al. (2001, 2002a) also found that the bias parameter depends on luminosity and spectral type: bright early type galaxies exhibit a relatively strong bias whereas faint late type galaxies show a slight anti–bias $b = 0.8$. Their luminosity dependence is quantified by:

$$\frac{b}{b_*} = 0.85 + 0.15\frac{L}{L_*}. \tag{62}$$

Another correlation function analysis was performed by Madgwick et al. (2003) who separated galaxies in the 2dF Galaxy Redshift Survey, according to spectral type. At small separations they found that passive galaxies exhibit much stronger real–space clustering than actively star–forming galaxies – the correlation functions have slopes of 1.93 and 1.5 respectively – and they found that the bias between the two classes declines as a function of scale. On scales larger than $10h^{-1}\mathrm{Mpc}$ they found no biassing between the two samples. There is thus not only evidence that biassing depends on spectral type, luminosity and morphology, but clear evidence of a scale

dependence in the biassing relation. Dekel and Lahav (1999) develop a formalism for dealing with these effects by separating the complexities of the biassing scheme into non–deterministic and non–linear parts. A non–linear biassing would reflect scale–dependence whereas local physical effects due to spectral type, luminosity etc. are quantified through a random stochastic factor. The relationship between the fields, in Fourier space, may be written:

$$\delta_{gal}(k) = b_L(k)\delta_{mass}(k) + \epsilon \tag{63}$$

where $b_L(k)$ reflects possible scale–dependence and ϵ is a random factor incorporating the local effects. This latter stochastic term motivates the definition of a correlation coefficient r_g between the two fields:

$$r_g = \frac{\langle \delta_g \delta_m \rangle}{\sqrt{\langle \delta_g \delta_g \rangle \langle \delta_m \delta_m \rangle}}. \tag{64}$$

This parameter is the subject of much discussion in Chapters 3 and 4.

1.11.4 Redshifts and Velocities

The crudest biassing scheme assumes a constant factor between the galaxy and mass fields at all wavelengths:

$$\delta_g = b\delta_m. \tag{65}$$

One indication that this scheme is too simplistic is that by definition $\delta_g \geq -1$ placing an unreasonable constraint on δ_m for $b > 1$. In Chapter 3 the more complicated biassing scheme of Dekel and Lahav (1999) is used which attempts to model the plethora of effects discussed in the previous section. Assuming equation (65) for now, the galaxy power spectrum can then be modelled in terms of the matter power spectrum by:

$$P_{gg}(k) = b^2 P_{mm}(k). \tag{66}$$

Although this equation follows from equation (65), the inverse is not true. In fact the original description of the biassing of Abell clusters by Kaiser (1984), was an attempt to model the bias introduced into the correlation function by the fact that Abell clusters only formed where the density enhancement was unusually large. This led to a biassing in the two point distribution as described by equation (66) but this does not automatically imply equation (65). From equation (42) it is clear that the peculiar velocity field does directly relate to the underlying matter.

The power spectrum of radial velocities may therefore be written:

$$P_{uu}(\mathbf{k}) = \mu^2 H_0^2 f(\Omega_m)^2 P_{mm}(k) \tag{67}$$

where $\mu = \cos\theta$ corresponds to the angle between $\mathbf{k}$ and $\hat{\mathbf{r}}$. The density fluctuations are modelled here as plane waves: $\delta_r(r) = \delta_k \cos(\mathbf{k}.\mathbf{r})$. In order to model both power spectra it is necessary to have a model for the underlying matter power spectrum $P_{mm}(k)$. This involves the use of the transfer function.

1.11.5 The transfer function

The primordial power spectrum is assumed to be a simple power law with k, which becomes modified during the expansion. This modification is characterized by the transfer function which takes into account the suppression of high k–modes by various damping processes. The transfer function is defined as:

$$T(k) \equiv \frac{\delta(k, z=0)}{\delta(k, z=\infty)} \frac{\delta(0, z=\infty)}{\delta(0, z=0)} \tag{68}$$

and the power spectrum becomes

$$P(k) = Q k^n T^2(k) \tag{69}$$

where Q is the amplitude. Most analyses adopt the scale invariant Zeldovich spectrum with $n = 1$. Poisson's equation gives: $\Phi(k) = -4\pi G\rho_0\delta(k)/k^2$ in phase space for a potential perturbation $\Phi(k)$. It is common to quote the power spectrum in a dimensionless way using the combination $k^3P(k)$. For the Zeldovich spectrum, the dimensionless power due to the perturbation Φ is thus independent of k – the potential field is a fractal – hence the term scale invariant. In fact a small deviation, $n = 0.97$, referred to as tilt, has been detected by WMAP (Spergel et al. (2003)). A small deviation is in fact predicted in inflationary scenarios. The primordial power law spectrum is damped by various processes and this damping is characterised by the transfer function.

Figure 5 compares three different transfer functions (taken from Eisenstein and Hu (1998)). The lines show the ratio of the transfer functions obtained to the correct zero baryon transfer function, using three different models for including baryons. The solid line corresponds to the

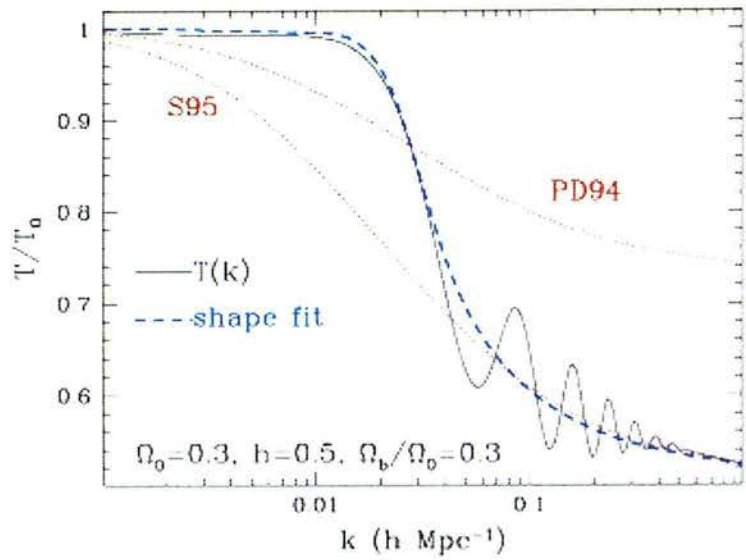

Figure 5: *Comparison of different transfer functions (taken from Eisenstein and Hu (1998)).*

high accuracy results of numerical codes such as the CMBfast code. The other lines correspond to the different fitting functions. PD94 is the transfer function of Peacock and Dodds (1994) and S95 is that of Sugiyama (1995), both of which use the zero-baryon transfer function of Bardeen et al. (1986) and account for baryons by adopting an effective shape parameter. The best fitting dashed line is the transfer function of Eisenstein and Hu which more accurately deals with the suppressing effects of baryons on the power spectrum .

The forms of these fitting functions are motivated by different damping processes. At low wavenumber the power still follows the primordial power law and all three transfer functions tend to one. The main cause of the suppression of power at high k is the Mészáros effect. The amount of suppression depends on the amount of time spent by a mode, within the particle horizon before matter radiation equality. The smallest wavelength modes enter the horizon first and so are the most damped. The break in the transfer function coresponds to a wavelength of the same scale as the horizon size at the time of matter radiation equality. Modes of higher wavelength are not damped by the Mészáros effect. The break scale in the transfer function thus allows the constraint of $\Omega_m h$ since this corresponds to the horizon size.

The nature of dark matter also affects the amount of damping. Relativistic dark matter particles as represented by HDM can free stream out of perturbations wiping them out. There is

therefore a damping of HDM perturbations which lie within the causal horizon before the particles become non-relativistic. Another important scale in the transfer function is therefore the horizon size at this time. A similar damping will occur for CDM but the particles will become non-relativistic sooner, and the corresponding scale will be smaller. A purely HDM density field would display more suppression than a purely CDM field. The existence of galaxies at $z \approx 6$ causes problems for HDM models.

Baryons have two effects on the transfer function – an oscillatory effect due to acoustic waves in the plasma before combination, and an overall suppression due to Silk damping (Silk (1968)). The transfer function of Eisenstein and Hu deals particularly well with the latter effect. Silk damping is the result of convection of the baryons caused by diffusion of photons in the plasma. The critical scale here is the mean free path for photons. Again modes with smaller wavelength will suffer greater damping and the overall effect is a further suppression of power on small scales.

1.11.6 Redshift Space Distortions

Two distortion effects arise in the observed galaxy distribution from the use of redshifts to predict distances. The observed redshift is the sum of two terms: the Hubble redshift due to the expansion and a Doppler shift due to the peculiar motions of the galaxies. The distortion is purely radial causing anisotropy in the redshift–space power spectrum. The distortions manifest themselves in two ways. On small scales the effect of virialised velocity fields is to stretch out the appearence of structures in redshift space, creating the so called Fingers of God effect. On larger scales, coherent infall into clusters causes a flattening of observed structures.

Models of the latter effect, based on the above results for the velocities induced by gravitational potential perturbations, allow the constraint of Ω_m from the distortion. A simplified model of the effect by Kaiser (1987a) uses the plane wave approximation. If s is a radial displacement in redshift space and r is the corresponding coordinate in real space then the two are linked by

$$s = r\left(1 + \frac{u(\mathbf{r}) - u(\mathbf{0})}{r}\right) \tag{70}$$

where $u(0)$ is the radial velocity of the observer. The volume elements transform as

$$d^3s = d^3r\left[1 + \frac{u(\mathbf{r}) - u(\mathbf{0})}{r}\right]^2\left[1 + \frac{du(\mathbf{r})}{dr}\right]. \tag{71}$$

Densities in real and redshift space are linked by

$$\frac{\rho_r}{\rho_s} = \frac{d^3s}{d^3r}\frac{\phi(\mathbf{s})}{\phi(\mathbf{r})} \tag{72}$$

where ϕ is the selection function. Taylor expanding the selection function in s–space about r, and using equation (71), gives to first order

$$\delta_s = \delta_r - \left[\frac{du}{dr} + \frac{u(\mathbf{r}) - u(\mathbf{0})}{r}\left(2 + \frac{d\ln\phi}{d\ln r}\right)\right]. \tag{73}$$

For a plane wave density perturbation, the Fourier mode

$$\delta_r(r) = \delta_k\cos(\mathbf{k}.\mathbf{r} + \theta) \tag{74}$$

can be substituted into equation (42) to give:

$$\frac{du}{dr} = -\mu^2\Omega_m^{0.6}H\delta_r(r) \tag{75}$$

where μ is the cosine of the angle between $\mathbf{k}$ and $\mathbf{r}$, and the growth rate of perturbations $f \approx \Omega_m^{0.6}$ using Peebles' approximation. Putting this into equation (73) and neglecting the last term which is negligible for a deep survey, gives

$$\delta_s(\mathbf{r}) = \delta_r(\mathbf{r})(1 + \Omega_m^{0.6}\mu^2). \tag{76}$$

The Fingers of God effect is usually modelled as an exponential smoothing of individual modes. This smoothing leads to the boosting of power along the line of sight. Figure 6 shows the counteracting effects of the distortions on a spherical cluster of galaxies. The galaxies are represented by dots and their peculiar velocities are shown by the arrows. The overdensity is spherical in real space. On large scales the inward pointing peculiar velocities give the observed structure a squashed appearance. The extreme of this, shown in the middle, is that the structure appears to have completely collapsed. On the smallest scales the peculiar velocities conspire to invert the positions of galaxies observed in redshift space. This effect combined with that of virialised incoherent velocity fields leads to the Fingers of God effect.

Although Kaiser's model is useful for small angular scales, where overdensities really can be considered plane waves, for wide field redshift surveys the correct analysis requires

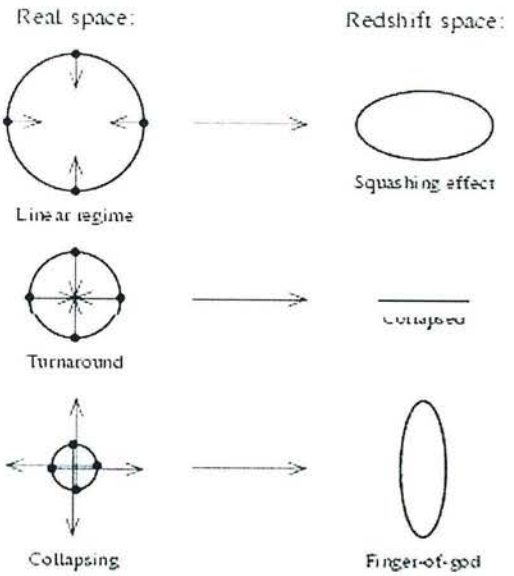

Figure 6: *Schematic illustration of redshift distortions (taken from Hamilton (1998)). See text for details.*

a description of the spherical nature of the survey. In Chapter 4 a more accurate method of modelling redshift space distortions is discussed in the context of an expansion in spherical harmonics.

1.12 Thesis Overview

In Chapter 2 the study of large scale structure through redshift and velocity surveys is reviewed. Different types of analysis of redshift surveys are discussed and some of the key results are outlined. Velocity surveys are also reviewed. Methods for obtaining peculiar velocities are discussed along with key results from the first velocity surveys to the present day. Finally Chapter 2 describes the 6dF Galaxy Survey, a key subject of later chapters.

Chapter 3 concerns an information analysis of the 6dFGS. The Fisher information matrix is used to predict the accuracies within which the 6dFGS can constrain cosmological parameters, assuming reasonable models for the shot noise of the survey. The Fisher matrix is used to optimise the survey design in terms of information content on the power spectrum. New results are then found for the combined Fisher matrix for two fields - the density and velocity fields.

Using these results the improvements on 6dFGS parameter constraints from the inclusion of peculiar velocities are found.

Chapter 4 outlines a method for the maximum likelihood analysis of the 6dFGS and other surveys. The work in this chapter builds on the spherical harmonics formalism of Heavens and Taylor (1995), extending it to the joint analysis of redshifts and velocities. First of all the method is introduced and simulations of the 6dFGS are constructed as a test bed. Data compression is needed and this is done optimally through Generalised Optimal Mode Analysis. The results of testing the software on the simulations are then presented and discussed. The method is found to accurately reproduce cosmological parameters input into the simulations, verifying both the formalism and the software.

2 Galaxy Surveys: Past, Present and Future

This chapter reviews the work done using redshift and peculiar velocity surveys from the first redshift surveys up to the present day. Section 2.1 discusses the history of redshift surveys, while Section 2.2 discusses key findings, in particular recent results from the 2dF and SDSS surveys. Section 2.3 introduces velocity field analysis beginning with a description of different techniques used to obtain redshift independent distances. Section 2.4 focuses on the measurement of bulk flows. There is then a discussion of the use of the various peculiar velocity catalogues for different analyses and an overview of the main results (Sections 2.5 and 2.6). Finally Sections 2.7 and 2.8 introduce the 6dF Galaxy survey describing the design of the survey and its motivations.

2.1 Galaxy Redshift Surveys

2.1.1 Optical Redshift Surveys

The very first redshift surveys in the mid 1970s were pencil beam surveys of nearby clusters which showed inhomogeneity in the galaxy distribution for the first time (e.g. Gregory and Thompson (1978)). The first wide angle survey was carried out by Sandage and Tammann (1981) with a survey which had a median radial velocity of 1500 kms^{-1} – too shallow to detect structure beyond the Local Supercluster. A greater depth was probed by the CfA Redshift survey (Huchra et al. (1983)) which confirmed the existence of large scale structure but was limited by sparse sampling.

A deeper survey still was completed by de Lapparent et al. (1986) but the thin slice geometry of the survey did not allow a complete 3 dimensional picture to emerge. The Perseus–Pisces region was surveyed in the HI Arecibo survey confirming the existence of large coherent structures. A southern hemisphere extension of the CfA survey – the Southern Sky Redshift Survey (SSRS, da Costa et al. (1988)) detected a thin structure called the Southern Wall. Extensions of both the CfA and SSRS with fainter magnitude limits and therefore greater sampling were performed. The CfA2 survey (Geller and Huchra (1989)) detected the Great Wall

– a thin 2–dimensional structure made up of galaxy clusters. The SSRS2 (da Costa et al. (1994)) confirmed the common occurrence in the southern hemisphere of voids and walls. The combined catalogues of both the CfA2 and SSRS2 surveys covered 30 % of the sky down to $cz \leq 15,000 \mathrm{kms}^{-1}$, allowing the analysis of scales up to 300 $h^{-1}\mathrm{Mpc}$. The Stromlo–APM survey (Loveday et al. (1992)) measured redshifts for 1800 galaxies drawn randomly at a rate of 1 in 20 from the APM galaxy survey. This survey probed a depth of $\sim 200h^{-1}\mathrm{Mpc}$ – a volume around five times the combined CfA2– SSRS2 surveys, but with much poorer sampling, at the expense of small scale information. The data were used to measure the luminosity function of galaxies on large scales as well as clustering statistics. In the mid–nineties the Optical Redshift Survey (Santiago et al. (1995)) measured redshifts for roughly 1300 optically selected galaxies to complete a sample compiled from the literature of over 8000 objects. This provided at the time the most detailed and homogeneous sampling of the large–scale galaxy distribution in the region $20^o \leq |b| \leq 30^o$ and $-17.5^o \leq \delta \leq -2.5^o$.

2.1.2 Infrared–selected samples

All optical surveys suffer from dust obscuration of galaxies by the plane of our own galaxy – defining the 'Zone of Avoidance' region on the sky. Selection in the infrared circumvents this issue and so surveys based on IRAS (Infrared Astronomical Satellite) galaxies such as the 1.2 Jy IRAS survey (Fisher et al. (1995a)), QDOT (Queen Mary and Westfield College, Durham, Oxford and Toronto, Saunders et al. (1991)) and the PSCz (Point Source redshift survey, Saunders et al. (1998)), were able to obtain a more uniform sky coverage. The PSCz contained 15,500 galaxies and covered more than 83% of the sky. One of the benefits of a uniform all sky survey is that it greatly simplifies statistical analyses because it removes problems associated with complicated window functions and edge effects.

2.1.3 Multi–object spectrographs

A great advance in redshift surveys came with the use of multi–object spectrographs on wide field telescopes. The Las Campanas redshift survey (LCRS, Shectman et al. (1996)) obtained redshifts for more than 25,000 R band selected galaxies in six strips covering 0.2 steradians.

The median redshift was $z \sim 0.1$ and so the volume covered was about five times larger than the combined CfA2–SRSS2. Other surveys with comparable depths included the Century survey (Geller et al. (1997)) and the ESO Project Slice (Vettolani et al. (1997)). All of these surveys presented a picture of a network of voids of roughly $\sim 5000\mathrm{kms}^{-1}$ diameter surrounded by large thin walls with no structure on larger scales. The surveys were however limited in the scales over which they could cover.

The current state of the art in wide field redshift surveys of the local Universe, are the recently completed 2 Degree Field Galaxy Redshift Survey (2dFGRS) (Colless et al. (1999)) and the Sloan Digital Sky Survey (SDSS) currently underway (first data release: Abazajian (2003)). The 2dFGRS consists of ~ 250000 galaxies selected from the extended APM survey, mainly contained within two strips of declination in the northern and southern galactic hemispheres. The redshift distribution has a median of 0.11 and so the 2dF galaxy survey has similar depth to LCRS. The SDSS when completed, will contain $\sim 10^6$ galaxies over about a quarter of the sky in two patches.

2.2 Analysis of Redshift surveys

2.2.1 The Luminosity function

The Luminosity Function (LF) $\Phi(L)$ gives the number density of galaxies with luminosity between L and $L + dL$. The shape and amplitude of this function, as well as its evolution with redshift are of fundamental importance in studies of the galaxy population. Added to this is the importance of the LF in estimating the selection function of any magnitude limited survey. This can be found using:

$$\phi(r) = \frac{\int_{4\pi r^2 f_{min}}^{\infty} \Phi(L)dL}{\int_{L_s}^{\infty} \Phi(L)dL}. \tag{77}$$

where f_{min} is the detectable flux limit and L_s is some limit below which the LF is small and poorly constrained. The LF is also used as a way of calibrating semi–analytical galaxy formation models. Because of its importance in all of these areas, the Luminosity Function has been the focus of much attention in redshift survey analysis. The local universe LF was

calculated from the CfA2 and SSRS2 surveys split into north and south subsamples. The shapes of all four were in good agreement but the normalization of the northern CfA2 survey was higher than the rest implying a galaxy density twice that of the other regions. The other three subsamples had consistent normalizations. Other more distant samples showed the same dichotomy – falling into either the high or low normalization groups but all with the same shape. Possible explanations for this included the existence of a large underdensity in the local Universe or rapid evolution of the blue Luminosity Function at low redshifts.

Other work has involved separating objects according to colour and morphology and calculating separate LFs for comparison. Analysis of the CfA2 and SRSS2 samples showed that early and late type galaxies have a similar flat LF whereas for irregular galaxies the slope is much steeper. Analysis of different colours showed an excess of blue galaxies at low magnitudes. Deeper surveys such as the Canada France Redshift survey have been used to study the evolution of the Luminosity Function with redshift. These showed that blue galaxies exhibited strong evolution while red galaxies exhibited very little.

More recently the 2dFGRS has allowed the determination of the Luminosity Function in the optical (Norberg et al. (2002)) and in the near–infrared (Cole et al. (2001)). Norberg et al. calculate the b_J band LF at zero redshift for 110500 galaxies. This forms the selection function for the 2dFGRS. Cole et al. select $\sim$ 17000 galaxies from the 2dFGRS according to their infrared magnitudes as catalogued in the Two Micron All Sky Survey (2MASS). This near–infrared Luminosity Function is intrinsically related to the mass distribution in galaxies.

2.2.2 The Power Spectrum

Calculation of the power spectrum or the correlation function provides a fundamental test of the gravitational instability paradigm, since the model predicts non–linear collapse on small scales and subsequent distortion of the matter power spectrum. Simple assumptions about the biassing of galaxies with respect to the dark matter allow the gravitational instability paradigm (and the biassing relation) to be tested. There is also information in these statistics about the clustering properties of different galaxies. Early on it was noticed that the power spectrum calculated from optically selected redshift surveys differed in amplitude from that calculated

using infrared selection. This was the first indication that the galaxy distribution is biassed with respect to the underlying mass, in a way which depends on galaxy type.

The power spectrum can also be used to place constraints on cosmological parameters by fitting the observed power to a model which takes into account the growth of perturbations since decoupling. A common parameter to constrain is the mass density parameter Ω_m but this is degenerate in the break scale of the power spectrum with the Hubble parameter, H_0. Sutherland et al. (1999) used the PSCz survey to constrain the shape parameter at $\Gamma = \Omega_m h = 0.25 \pm 0.04$. Percival et al. (2001) measured the power spectrum of the 2dFGRS on scales up to $300\mathrm{h}^{-1}$ Mpc. The shape was found to be consistent with a CDM model with shape parameter $\Gamma = \Omega_m h = 0.20 \pm 0.03$. Assuming a present day Hubble parameter of $H_0 = 70\mathrm{kms}^{-1}\mathrm{Mpc}^{-1}$ this gives the density parameter as $\Omega_m \approx 0.3$. The power spectrum was also used by Percival et al. to detect baryonic damping, giving $\omega_b = \Omega_b/\Omega_m = 0.15 \pm 0.07$.

Tegmark (2003b) calculated the real space power spectrum for 205,443 galaxies from the SDSS. They measured the shape parameter as $\Gamma = 0.201 \pm 0.017$ and the rms galaxy fluctuation within spheres of $8h^{-1}\mathrm{Mpc}$, giving $b\sigma_8 = 0.89 \pm 0.02$ for L_* galaxies.

Greater constraints on cosmology were possible by combining the 2dFGRS power spectrum with CMB data. Efstathiou et al. (2002) and Percival et al. (2002) found that the Universe is flat ($\Omega_k = 0 \pm 0.05$) with a total matter density $\Omega_m = 0.31 \pm 0.06$ and contribution from the cosmological constant of $\Omega_\Lambda = 0.69 \pm 0.06$. This is consistent with the independent measurement from supernovae (e.g Perlmutter et al. (1999)). The 2dFGRS and CMB datasets also gave cold dark matter and baryon densities of $\omega_c = \Omega_c h^2 = 0.12 \pm 0.01$ and $\omega_b = \Omega_b h^2 = 0.022 \pm 0.002$ respectively (Percival et al. (2002)). They also gave an estimate of the Hubble parameter of $H_0 = 67 \pm 5\mathrm{kms}^{-1}\mathrm{Mpc}^{-1}$. The estimated baryon fraction agrees with the predictions of Big Bang nucleosynthesis. An impressive consistency is found when comparing the results of the 2dFGRS+CMB results with the analysis of Tegmark (2003a) who combined their power spectrum measured from SDSS with the WMAP results. Their constraint of the matter density tightened to $\Omega_m = 0.25 \pm 0.1$ when they included the WMAP data. The consensus from all of these surveys is a flat Lambda–CDM model.

2.2.3 The Peculiar Velocity Field from Redshift Surveys

Peculiar velocities can be obtained through redshift independent measurements of galaxy distance. This is discussed in the next section, but redshift surveys on their own can be used to detect the peculiar velocity field by considering the anisotropy it causes in statistical measures such as the power spectrum or the correlation function (Kaiser (1987a)). Taylor et al. (2001a) use an optimal spherical harmonics approach to constrain the distortion parameter for IRAS galaxies in the PSCz. They find $\beta = 0.39^{+0.14}_{-0.12}$. The correlation function which is just the real space counterpart to the power spectrum, has also been the subject of much work. Consistency has been found between surveys of different sizes and different regions. Due to redshift space distortions however, the correlation function becomes compressed in the radial direction. This anisotropy allows the distortion parameter β to be measured. The 1.2 Jy IRAS survey gave estimates of $\beta = 0.45$ (Fisher et al. (1994a)). Due to the limited sampling of these surveys, this estimate suffered from cosmic variance and large systematic errors. Using the 2dFGRS Peacock et al. (2001) obtained $\beta = 0.54 \pm 0.09$ for galaxies with similar luminosity to the Milky Way. Since the completion of the 2dFGRS Hawkins et al. (2003) have performed a correlation function analysis on a larger set of $\sim$ 220000 redshifts. They find $\beta = 0.47 \pm 0.08$ for L^* galaxies: assuming a simple biassing of $b = 1$ this implies a present day density parameter of $\Omega_m \sim 0.3$ consistent with other 2dFGRS analyses. They also treat the redshift distortion due to random peculiar velocities finding, consistent with other analyses, that this is best fit by an exponential form for the pairwise velocity dispersion.

2.2.4 Galaxy Biassing

An important characteristic of the clustering patterns is the way in which galaxies of different types exhibit different clustering strengths. Not only is this important for the understanding of galaxy formation, it is also helpful in understanding the nature of the biassing between dark matter and galaxies. Work on the SSRS2 (Benoist et al. (1996), Willmer et al. (1998)) has shown scale independent, luminosity bias – i.e. more luminous galaxies exhibit greater clustering strengths. The SSRS2 catalogue also shows a relative bias between elliptical galaxies and spirals. This bias is found to be scale dependent, the difference being greatest on small scales. Differences are also found between galaxies of different colour and between optical and

IRAS galaxies, all of which suggests complications in the way the complete galaxy distribution relates to the dark matter. Analysis of the PSCz catalogue shows an IRAS bias parameter of $b = 1.22 \pm 0.28$ (Taylor et al. (2001a)).

Further work on galaxy biassing has been done using the 2dFGRS. Lahav et al. (2002) combined the 2dF power with that of the CMB to obtain the biassing of L_* galaxies. They find that $b^* = 0.96 \pm 0.08$ implying that the distribution of L_* galaxies perfectly traces the underlying mass. Verde et al. (2002) perform a higher order statistical analysis of the 2dFGRS on its own and find a similar result. However Norberg et al. [2001, 2002a] find that the bias parameter exhibits a linear dependence on luminosity as well as changing with spectral type, with the extremes of $b = 1.5$ for bright early type galaxies and $b = 0.8$ for faint late type galaxies. Madgwick et al. (2003) compare the clustering properties of galaxies of different spectral type in the 2dFGRS. They calculate the redshift space correlation function for galaxies separated according to star formation rate. They find a relative bias between the two samples (with active and passive evolution) which decreases as a function of scale. They do however find similar degrees of redshift space distortion between the two classes with $\beta_{active} = 0.49 \pm 0.13$ and $\beta_{passive} = 0.48 \pm 0.14$.

2.2.5 Wiener reconstruction of the density field

Wiener filtering – as discussed for example in Rybicki and Press (1992) – is a way of suppressing shot noise and dealing with selection and incompleteness effects, in smoothed density fields reconstructed from galaxy surveys. The method can be equally applied to reconstruction of the velocity field. Wiener filtering consists of making a linear combination of a data set, optimally chosen to minimise the variance between the signal obtained from the data and a hypothetical 'true' signal predicted from some fiducial model. In this sense, as with the Karhunen Loeve eigenvalue compression discussed in Chapter 4, the Wiener filter requires a priori assumptions about the cosmological model. The Wiener filter is

$$\mathbf{W} = \mathbf{C}[\mathbf{C} + \mathbf{N}]^{-1} \tag{78}$$

where $\mathbf{C}$ is the modelled covariance of the signal and $\mathbf{C} + \mathbf{N}$ is the covariance of the data which includes noise. The filtered data is just the original data vector multiplied by the Wiener

filter matrix. There is a wealth of information in the literature regarding the application of the Wiener filter to studies of large scale structure. The formalism for applying the filter to noisy, sparse and incomplete data is introduced in Zaroubi et al. (1995). Erdogdu et al. (2004) apply the Wiener reconstruction method to the 2dFGS in the search for superclusters and voids. Their most surprising result is found when their data are used in conjunction with the 2MASS and Las Campanas Survey (Frith et al. (2003)): they find evidence for a large contiguous void stretching from north to south, creating power on scales at which homogeneity is thought to set in. The 2dFGS is fundamentally limited on its own however because of its relatively small sky coverage. The 6dFGS should allow Wiener reconstruction of both the density and velocity fields providing amongst other things, a means of constraining the distortion parameter β.

2.3 Different distance indicators

The most direct way to measure the peculiar velocity of a galaxy is to obtain a redshift–independent measure of its distance. This can then be subtracted from the redshift to give the radial peculiar velocity. To obtain redshift independent distances it is necessary to use a correlation between an observable distance–dependent quantity and a distance–independent quantity. These distance indicators can be used to infer peculiar velocities or they can be used to calibrate the Hubble relation. The use of supernovae as standard candles is one of the most notable examples of the latter. Although supernovae have proved supremely useful in detecting changes in the Hubble parameter, they are of limited use in the detection of large scale flows because of their rarity and unpredictability. In this section a few methods are briefly discussed which can be used for large surveys of the local Universe.

2.3.1 The Tully Fisher relation

The first peculiar velocity surveys involved spiral galaxies. The Tully Fisher relation is an empirically observed correlation between the rotation velocity of a spiral galaxy and its luminosity (Tully and Fisher (1977)). The rotation curves of spiral galaxies are observed to flatten off beyond the centre. Because of this, it is reasonable to assign to such galaxies a single characteristic rotation velocity. It is also intuitively reasonable for a correlation to exist between

luminosity and rotation velocity since both must be linked somehow to the mass of the galaxy. A theoretical derivation of the Tully Fisher relation has not however been achieved. The systematic uncertainties surrounding its derivation dominate the exact form of the relation and so an exact physical understanding is not strictly necessary for use as a distance indicator. The underlying physical correlation is

$$L(v_{rot}) \propto v_{rot}^{\alpha} \tag{79}$$

where α is to be empirically determined. In terms of apparent magnitude the Tully Fisher relation may therefore be written:

$$m(\nu) = \mu(r) + A - b\nu \tag{80}$$

where $\nu \equiv \log(2v_{rot}) - 2.5$ is a convenient measure of the rotation velocity and $\mu(r) = 5 \log r$ is the distance modulus. Once the Tully Fisher relation is calibrated and A and b have been found, it just remains to observe the rotation velocity and the apparent magnitude in order to find the distance. The Tully Fisher relation was originally applied to HI 21 cm profiles. The images were not resolved but the width of the line and the inclination of the galaxy (obtained from its ellipticity) allowed the rotation velocity to be estimated. Later work involved optical observations and used the Hα line (e.g. Dressler and Faber (1990)). Use has also been made of an infra red Tully Fisher relation (Aaronson et al. (1982)).

2.3.2 The $D_n - \sigma$ relation

The equivalent distance indicator for elliptical galaxies is the Faber Jackson relation (Faber and Jackson (1976)). This uses an observed correlation between velocity dispersion and luminosity. The velocity dispersion in question is the scatter in the radial velocities of stars in the centres of the galaxies. The correlation takes the form:

$$L \propto \sigma_e^{\alpha}. \tag{81}$$

The scatter in this relation is roughly twice that of the Tully Fisher relation. An improvement can be made by introducing a third parameter, as discovered by Djorgovski and Davis (1987) and Dressler et al. (1987b). The data in this three dimensional parameter space is found to occupy a *Fundamental Plane*. The correlation is between an effective radius R_e, average

stellar velocity dispersion σ_e and surface brightness I_e within that radius:

$$R_e \propto \sigma_e^{\alpha} I_e^{-\beta}. \tag{82}$$

Finally it is possible to define a single parameter to encompass the surface brightness and radius terms giving the correlation studied by Lynden-Bell et al. (1988):

$$D_n \propto \sigma_e^{\gamma} \tag{83}$$

where D_n is the diameter within which the surface brightness averages to a certain value. The '7 Samurai' group found that the accuracies of distances inferred from the $D_n - \sigma$ relation were twice as accurate as the Faber–Jackson relation and comparable with those found using the Tully Fisher relation with spirals. Using these distances they were able to infer previously unknown significant peculiar velocities on large scales.

The use of both Tully Fisher and $D_n - \sigma$ relies on their correlations being universally true. Early on there had been concerns about potential differences in the relations when calibrated using galaxy clusters and when using field galaxies. Kolatt and Dekel (1994) however, compared velocity fields inferred from spiral galaxies (TF) and from elliptical galaxies ($D_n - \sigma$) and found no systematic differences.

2.3.3 Surface Brightness Fluctuations

Another way of obtaining distances to elliptical galaxies is to use the Surface Brightness Fluctuation technique on pixellated images. The Poisson fluctuations of the number of stars observed in a pixel depend on the average number encompassed by the pixel which in turn depends on the distance of the galaxy. There is therefore a correlation between surface brightness fluctuations and distance. The method is very accurate for galaxies with redshifts $z \leq 0.1$ which are close enough to be resolved. For these galaxies the distances have $\sim 5\%$ accuracy, but for higher redshifts the techinique is hindered by seeing. The $D_n - \sigma$ method has the advantage of being useful at the greater depths probed by large galaxy surveys such as the 6dFGS.

2.3.4 Brightest Cluster galaxies

At the other end of the distance scale, the most luminous objects in the centres of rich clusters – or Brightest Cluster Galaxies – can be used. Hoessel (1980) found that the luminosity of these objects correlates with the shape of the galaxy's luminosity profile. Specifically, a luminosity shape parameter is defined as

$$\alpha = \left. \frac{d \log L}{d \log r} \right|_{r_m} \tag{84}$$

and this is found to correlate with L_m defined as the total luminosity within a radius $r_m = 10h^{-1}\mathrm{kpc}$. This method was used by Lauer and Postman (1994) who obtained distance errors of $\sim 16\%$. The Lauer and Postman result was rather controversial and is discussed below.

2.3.5 Satistical biases in peculiar velocity analysis

Having obtained distance indicators and redshifts using one of the above methods there is then a further choice of how to analyse the data. It is convenient to label different methods according to the way they assign the positions of galaxies. Faber and Burstein (1988) refer to analyses which choose the distance indicator as the estimate of distance as Method I analyses. When redshifts are used as the true positions the analysis is referred to as Method II. Within these definitions there is another methodological division. A 'Forward' analysis uses the distance independent quantity to predict the distance dependent quantity. So for the $D_n - \sigma$ relation a forward analysis would use the velocity dispersion to predict the diameter D_n. The 6dF velocity survey will have this approach: using the absolute diameter to predict distances from the apparent diameter. An 'inverse' method conversely, uses the distance dependent quantity to predict the distance independent quantity.

The type of analysis used is important because different statistical biases come into play. Selection bias affects method II analyses. A common example is that a flux limited survey will have high redshift galaxies biassed towards high magnitudes. Thus for example, if the $D_n - \sigma$ relation is calibrated using this data the smallest diameter galaxies will be missing from the sample, and the inferred correlation will be incorrect. Method I analyses on the other hand, suffer from Malmquist bias. Distinct from selection bias, Malmquist bias arises because distance indicators are not unbiassed estimates of true position. Redshift positions on average give

unbiassed estimates of true position, even though they are systematic deviations, but distance indicators are prejudiced towards the positions along the line of sight where their probability distribution peaks. In other words the biassing effect in distance indicator information depends on the true line of sight galaxy distribution which is in general unknown. The POTENT analysis (see below) is a Method I analysis and so suffers from Malmquist bias.

2.4 Measurements of Bulk Flows

2.4.1 Motion of the local group

Until 1976 it was assumed that the local group was at rest with respect to the expansion of the Universe. This assumption was altered by the discovery of the CMB dipole which can only reasonably be explained by a bulk motion of the local group rest frame with respect to the microwave background. Between 1976 and 1986 there were numerous attempts using the Tully Fisher and Faber Jackson relations to detect these bulk motions. Bulk flows are most commonly measured by a least squares fit of model predictions $\mathbf{V}_b.\mathbf{n}$ to each velocity, where $\mathbf{n}$ is a unit vector in the direction of each galaxy and $\mathbf{V}_b$ is the single bulk flow vector to be constrained. The result of Rubin et al. (1976) was a surprising flow orthogonal to the CMB dipole direction. Many more bulk flow detections were made mostly in the direction of the Virgo cluster. The explanation given to the CMB dipole flow, not aligned with Virgo infall, was attraction towards the Hydra–Centaurus supercluster. The Hydra–Centaurus supercluster is at a relatively local redshift of 0.01.

2.4.2 Measuring the dipole in the galaxy distribution

The peculiar gravity induced by the mass of a galaxy distribution may be found without the need for calculating redshifts as is explained in the following heuristic argument. Under the idealized assumptions that the galaxies have a constant mass to luminosity ratio and that the sample is volume limited, the peculiar gravity is obtained from:

$$\mathbf{g} = G \sum_i \frac{m_i}{r_i^2} \hat{\mathbf{r}}_\mathbf{i} \tag{85}$$

or,

$$\mathbf{g} \propto \sum_i f_i \hat{\mathbf{r}}_\mathbf{i} \tag{86}$$

where the sum is over galaxies with mass m_i, flux f_i and unit vectors $\hat{\mathbf{r}}_\mathbf{i}$. Meiksin and Davis (1986) and Yahil et al. (1986) both use an advanced version of this technique to obtain the peculiar velocity dipole from IRAS galaxies. These parallel approaches yield impressively consistent results which are proportional to the dipole anisotropy in the microwave background. Lahav et al. (1988) calculate the peculiar acceleration of the Local Group in this way and use it to estimate Ω_m, using the relation between peculiar velocity and peculiar gravity as described for example by equation (57). They find a high value of Ω_m from their IRAS sample and a lower value for their optical sample – another indication of galaxy biassing. A redshift survey follow–up to the work on IRAS galaxies was performed by Strauss et al. (1992) confirming the direction of the acceleration of the Local Group and that the CMB dipole is indeed caused by the motion of the Local Group induced by peculiar gravity.

2.4.3 The Great Attractor – 7 Samurai group

The accepted view that deviations from the Hubble flow died off after $cz \approx 500\mathrm{kms}^{-1}$ was challenged by the discoveries of the '7 Samurai' group (7S Dressler et al. (1987a)) who used the $D_n - \sigma$ relation to find a large scale bulk flow with respect to the CMB. They found that the Hydra–Centaurus galaxies also moved as part of this flow and so could not be causing it themselves. The 7S group reinterpreted their data in Lynden-Bell et al. (1988) as motion towards a 'Great Attractor' – a mass of $\sim 5.4 \times 10^{16} M_\odot$ at a distance of $4350 \pm 350\mathrm{kms}^{-1}$. The model actually assumed a point mass for the Great Attractor. Further modelling, this time without assuming results for the Virgo flow, inferred that previous calculations of Virgo infall had been overestimates because the presence of the Great Attractor had not been taken into account.

There soon became evidence of bulk flows on even larger scales with the discovery by Willick (1990) of a bulk motion of the Perseus–Pisces supercluster region on the other side of the sky from the Great Attractor. The flow was observed to be in the general direction of the Great Attractor. Bulk flow of the Great Attractor itself was observed by Mathewson et al. (1992) however subsequent reanalyses have suggested that this result was due to selection or

Malmquist bias or perhaps even an incorrect zeropoint calibration of the Tully Fisher relation.

2.4.4 The Lauer– Postman Result

A detection of bulk flows on much larger scales was made by Lauer and Postman (1994). Using the Brightest Cluster Galaxy $L-\alpha$ relation, they determined a mean velocity vector for a sample of 119 Abell clusters with redshifts $cz < 15,000\mathrm{kms}^{-1}$. They defined an Abell Cluster Inertial Frame and found that it moved relative to the Local Group in a nearly orthogonal direction to the CMB dipole. The detected flow was a net peculiar velocity of $689 \pm 178\mathrm{kms}^{-1}$ in the direction $l = 343^o$, $b = +52^o$. A flow of this size on such a large scale went against the consensus of opinion, which favoured a quiescent velocity field beyond $cz \sim 6,000\mathrm{kms}^{-1}$. Previous observations had found convergence of the field on these scales. The result also proved troublesome for popular theories of structure growth. N–body simulations of the Lauer and Postman sample, using six different scenarios for structure formation were performed by Strauss et al. (1995). These included the standard cold and hot dark matter. The simulations found the observed velocity vector to be highly unlikely – only $2-5\%$ of realizations matched the Lauer and Postman result, using popular models. Subsequent re–analysis of the data has shown that statistical biasses may account for this anomolous result. Confirmation of this can only now come with new data from future surveys such as the 6dFGS, which will also provide peculiar velocities out to $cz < 15,000\mathrm{kms}^{-1}$.

2.5 Other Analyses of Peculiar Velocities

There are several other ways to analyse peculiar velocity fields beyond simply calculating bulk flow vectors. The large uncertainties make individual velocity measurements meaningless but they can still be used to obtain statistical measures of the velocity field. These measures can in turn be used to constrain cosmological parameters. One such statistical measure is the velocity power spectrum which can be used to constrain the amplitude of mass clustering. Another option is to combine the velocity power spectrum with the galaxy power spectrum as discussed in Chapter 4 where a new method involving spherical harmonics is discussed. Historically research has involved use of the velocity correlation function. The scale on which this falls

to zero is a measure of the coherence length of the velocity field. Gorski et al. (1989) studied samples of spirals and ellipticals and found that the correlation function dropped to zero at separations of $2000\mathrm{kms}^{-1}$ fitting with the predictions of CDM structure formation.

Another useful statistic is the Cosmic Mach Number defined as the ratio of the small scale velocity dispersion to the large scale bulk flow. This gives a measure of the shape of the power spectrum, independent of the amplitude. Results of Monte Carlo simulations assuming CDM strongly rule out the observed Mach numbers in Strauss et al. (1993). This is a confirmation of the observed coldness of the velocty field on small scales. Other dark matter models fit the observed Mach number more closely.

2.5.1 POTENT

An important technique developed to make use of all of the information available in a velocity survey is POTENT (Dekel et al. (1993)). It assumes that the velocity field possesses no vorticity, as is reasonable in the local universe since vorticity decays with the expansion of the universe. With no vorticity the three dimensional velocity field can be written as the derivative of a potential:

$$\mathbf{v}(\mathbf{r}) = -\nabla\Phi. \tag{87}$$

This potential can be obtained by integrating the observed radial velocity field:

$$\Phi(\mathbf{r}) = -\int_0^r u(r', \theta, \phi)dr'. \tag{88}$$

To obtain $u(r)$ in the first place requires the smoothing of observed discrete velocities. This allows smoothing over small–scale vorticity and non–linear effects and gives an averaging of the velocity field – essential considering the large errors on individual velocities. The POTENT technique has been applied to the Mark II and Mark III catalogues – compilations of peculiar velocity samples where the raw data has been reanalysed to create a single homegeneous sample. The POTENT method has been used to create maps of the velocity field picking out the Great Attractor. The technique has also been used to compare the velocity fields of ellipticals and spirals separately, and to constrain the amplitude of mass fluctuations by fitting the velocity field to a power spectrum (Seljak and Bertschinger (1994)). The reconstructed velocity field can also be analysed in other ways to constrain values of Ω_m or the function $f(\Omega_m)$. In the

non–linear regime, second order perturbation theory links the skewness of the velocity field to $f(\Omega_m)$. The skewness however, is strongly effected by the errors and biases in the POTENT technique and so needs to be tested properly with Monte Carlo simulations. Another means of constraining $f(\Omega_m)$ is by making use of voids in the reconstructed mass density field. The density perturbation δ has a lower limit of -1 corresponding to the absence of matter. Since

$$\nabla.\mathbf{v} = -f(\Omega_m)\delta \tag{89}$$

this places an upper limit on the velocity divergence dependent on Ω_m. Dekel and Rees (1994) have applied this to a void in the POTENT reconstruction of the Mark III catalogue and find that $\Omega_m > 0.3$ with a certainty of 2.4σ. Again this method is subject to large uncertainties – this time caused by a reliance on voids: the areas in velocity fields with the least data.

2.6 Recent and Future analysis of cosmic flows

2.6.1 Bulk Flows

At the time of the Lauer and Postman result and the few years proceeding it, the field of cosmic flows was in a state of disarray. No cosmological models were consistent with the LP result, and other studies of large scale bulk flows found motions of similar amplitude but in different directions (Hudson et al. (2000), Willick (1999)). Since then however, important new data sets have emerged, as has an improved understanding of the complexities of biassing. Courteau and Dekel (2001) review the recent results and present a compilation which is reproduced in Figure 7.

They compare the measured bulk flows within a sphere of radius R with that predicted by a ΛCDM model. The solid line gives the prediction and the dashed lines represent 90% of the cosmic scatter in the Maxwellian distribution of V, when sampling only one sphere. The data from each survey has been rather crudely translated into the bulk flow of a 'top hat' sphere. It is also worth noting that this plot gives no indication of the directions of the bulk flow measured. The agreement however, between most of the data sets is promising. There are three results noticeably outside the 90% boundaries. The BCG result is that of Lauer and Postman (1994) and its deviation from the cosmological consensus is clear to see. The LP10 result is from a

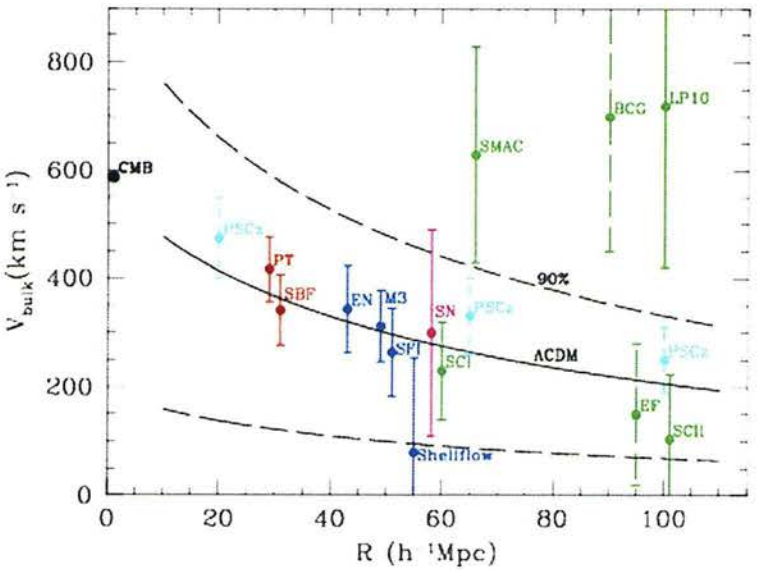

Figure 7: *Comparison of different bulk flow measurements with the ΛCDM model (from Courteau and Dekel (2001)). The labels are explained in the text.*

Tully Fisher survey of galaxies in 15 Abell clusters in the redshift range $9000 - 12000 \mathrm{kms}^{-1}$ (Willick (1999)). This result is very noisy and so should be treated with caution. Although the amplitude of the motion is in agreement with Lauer and Postman, the direction is different. The third outstanding bulk flow is from the SMAC (Streaming Motions of Abell Clusters) project – a Fundamental Plane study of ~ 700 early type galaxies in 56 local rich clusters (Hudson et al. (2000)). Other surveys shown find no large amplitude flows with respect to the CMB on the scales of the BGC and LP10 surveys. The SBF survey (Tonry et al. (2000)) uses the Surface Brightness Fluctuation technique on 300 early type galaxies and finds its survey volume of $R < 3000 \mathrm{kms}^{-1}$ to be at rest with respect to the CMB. EF labels the results of the EFAR project (Colless et al. (2000)) using peculiar velocities from the Fundamental Plane relation, for 85 clusters between 6000 and 15000 kms^{-1}. Their bulk flow does not support either the SMAC or LP10 results. ENEAR (Wegner et al. (2000)) is an all sky survey of nearby early type galaxies using the Fundamental Plane. Again the bulk flow is within cosmological predictions. PT is the result of the POTENT analysis applied to the Mark III catalogue and SN marks the results from type Ia supernovae (Riess (2000)). The Shellflow survey is an all sky Tully Fisher survey of galaxies between 4000 and 7500 kms^{-1}. Although slightly beyond the lower 90% boundary from ΛCDM models, the shellflow survey does find convergence to the Hubble flow

on large scales and so refutes the findings of BCG, LP10 and SMAC. The SFI, SCI and SCII are three samples of spiral galaxies with Tully Fisher distances. Again they confirm convergence of the velocity field on scales predicted by gravitational instability.

Some of the data plotted here are not from peculiar velocity surveys but derived from other data sets. The CMB point is just from the dipole in the CMB radiation, and the PSCz results are velocities inferred from the PSCz density field.

2.6.2 Velocity Power Spectrum

Improvements have also been made in the measurement of Ω_m from the power spectrum of peculiar velocities. Silberman et al. (2001) correct the form of the fitted power spectrum for non–linear effects, which cause the power on small scales to be larger than predicted in linear theory, and also cause suppression of the overall power on larger scales. They use high resolution mock catalogues of the Mark III and SFI catalogues to calibrate the correction to the power spectrum. The constraint on the density from the Mark III data changes from $\Omega_m = 0.56 \pm 0.04$ by fitting to a linear power spectrum, to $\Omega_m = 0.32 \pm 0.06$ using the improved model. Similarly results from SFI change from $\Omega_m = 0.51 \pm 0.05$ to $\Omega_m = 0.37 \pm 0.09$.

Another use of the velocity power spectrum is in its comparison with the galaxy power spectrum, again in order to obtain Ω_m. The problem with this is that it requires assumptions to be made concerning galaxy biassing. Uncertainties in the biassing relation also affect the construction of the mock catalogues, from which the non–linear effects in the velocity power spectrum are calibrated in the analysis above. Because of this, no velocity power spectrum analysis will be accurate until the intricacies of galaxy biassing are sorted out.

2.6.3 Future Velocity Surveys

Future data sets from ongoing surveys will improve on previous sample sizes by a factor of four. As well as 6dFGS there is the NOAO Fundamental Plane Survey (NFP), SNFactory and Warpfire. NFP will survey some 4000 early type galaxies, in 100 X–ray selected clusters over the whole sky and is expected to achieve $< 120 \mathrm{kms}^{-1}$ precision on each component

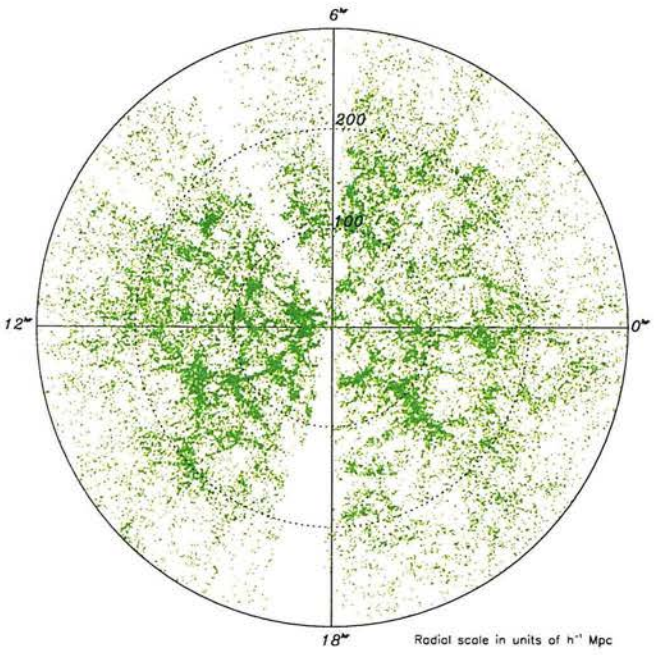

Figure 8: *Redshift Distribution from the 6dFGS (Colless et al. (2004))*

of the bulk flow vector. SNfactory is a supernovae project aimed primarily at measuring the equation of state of dark energy; supernovae have the advantage of unrivalled precision as a distance indicator, albeit that measurements are made at a rate of only a few hundred per year. Warpfire is the ongoing use of Brightest Cluster Galaxies by Lauer and Postman, which is being increased from the volume out to 15000kms^{-1} to a frame out to 24000kms^{-1}.

2.7 The 6dF Galaxy Survey

The Six Degree Field Galaxy Survey is the first combined redshift and peculiar velocity survey, using a new, purpose built multi object spectrograph on the UK Schmidt telescope (Parker et al. (1997); Wakamatsu et al. (2003)) . It currently has half of its target of 167000 redshifts and the whole survey is due for completion in mid–2005. Figure 8 shows the distribution to date, with extremely detailed structure visible. The redshift survey is drawn mainly from the 2MASS extended source catalogue (XSC), and so selection is performed in the near infrared. The velocity survey will use a subset of about 15000 early type galaxies from this sample. The velocities will be found using the $D_n - \sigma$ relation. The survey will cover most of the southern sky excluding galactic latitudes below $|b| < 10^o$. This corresponds to a total solid

angle of 17000deg^2. The redshift survey is magnitude limited down to a total magnitude of $K_{tot} < 12.75$. Total magnitudes are estimated using the isophotal magnitudes and surface brightness profiles found in the 2MASS XSC (Jarrett et al. (2000)). Other smaller samples are also used, drawn from the 2MASS and SuperCosmos catalogues. The median redshift is expected to be $z_{median} = 0.055$ or 15000kms^{-1}. There are also some additional target samples included for more specific scientific objectives.

2.7.1 Advantages of the Redshift Survey

6dFGS has several advantages over other surveys. The near infrared selection minimises dust obscuration both in the observed galaxy itself and in our own galaxy. The former means that measured mass to light ratios of the observed galaxies will be unaffected by their orientation to the line of sight. The latter allows the survey to be performed at relatively low galactic latitudes, despite dust in the line of sight near the galactic plane. This has the double advantage of providing a larger window than the 2dF survey, simplifying statistical analysis, and revealing previously hidden clusters. Near infrared selection also yields a better estimate of the galaxies' mass by picking out the old stellar population of the observed galaxies rather than being biassed by bursts of recent star formation.

2.7.2 Combination with the velocity survey

The velocity survey uses the same observations as the redshift survey and so together they comprise a homogeneous sample. The isophotal diameter D_n is calculated from the 2MASS images, whereas the velocity dispersion σ is found from the 6dF spectra. The dominant errors in the measurement of distances come from errors in the velocity dispersions and not from D_n. The errors in σ should be around 10% whereas the errors in D_n should be $\sim$ 0.03dex. The need for good resolution to calculate σ, means that velocities can only be calculated for galaxies within 15000kms^{-1}. The only other distinction between the selection criteria for redshifts and those for velocities, is that use of the $D_n - \sigma$ relation requires early type galaxies with high S/N spectra. Another advantage of the near infrared selection is that it is better for finding early type galaxies than optical selection which preferentially selects late types. There is no

possibility of using the Tully Fisher relation for the 6dF velocity survey, as this requires spatial information which is not available (Lachlan Campbell, private communication).

2.7.3 Clustering properties of the 2MASS galaxies

The near infrared selection yields a large number of elliptical galaxies which is important for the velocity survey, however this does affect the clustering statistics. Elliptical galaxies tend to be more strongly clustered that their spiral counterparts and are more likely to be found in galaxy clusters rather than in the 'field'. One expects therefore to measure a higher amplitude for the power spectrum of elliptical galaxies than for spirals.

2.8 Scientific Goals of the 6dFGS

The first goal of the 6dFGS is to measure the luminosity function of near infrared selected galaxies. This is strongly correlated with the stellar mass function of collapsed structures. 6dFGS will also be used to analyse the variation of the LF with spectral type, furthering the studies discussed in Section 2.2.1.

Other aims include mapping the local galaxy distribution, particularly that near the plane of the galaxy. Use will be made of the Wiener filtering method discussed in Section 2.2.5, for reconstructing the density field. The density power spectrum can be calculated to scales comparable with those of 2dFGRS and SDSS. Chapter 3 discusses the constraints on cosmological parameters that will be achievable from the 6dFGRS. The strength of clustering on different scales will be used to test the scale dependence of biassing. Again it is because of the near–infrared selection that the observed clustering should be close to the mass distribution.

The velocity survey will allow detailed mapping of the local peculiar velocity and density fields over half the local volume to $\sim 15,000 \mathrm{kms}^{-1}$. Again the Wiener reconstruction method can be used to counter the effects of shot noise and incompleteness. The combination of the velocity field power spectrum and galaxy power spectrum can be used to constrain β as in Section 2.6.2. The high quality spectra for the sample of early type galaxies in the velocity

survey, can also be used to study the ages, metallicities and star formation histories of early type galaxies over a range of masses and environments.

Overall the 6dFGS should provide an unprecedented opportunity to compare the galaxy clustering and velocity fields from a single homogeneous data set covering the entire southern sky. In the next chapter, predictions are made using the Fisher information matrix, about how accurately the surveys can be used to constrain cosmological models.

3 Fisher information analysis of galaxy redshift and peculiar velocity surveys

The work in this chapter was done in collaboration with A. N. Taylor and published in Burkey and Taylor (2003).

3.1 Introduction

In recent years the emergence of large data sets from galaxy surveys and CMB probes has greatly improved the constraint of cosmological parameters. In particular the first data release from WMAP combined with information from the 2dF Galaxy Redshift Survey and Lyman α forest data has allowed many cosmological parameters to be constrained to unprecedented accuracy (Spergel et al. (2003)). The Sloan Digital Sky Survey (SDSS) when complete, will constitute a fourfold increase in redshifts over 2dF. It is the aim of this chapter, in the context of the present data boom, to predict the unique advantages of the 6dFGS over its contempories.

This chapter uses an information–theory analysis of galaxy redshift surveys and peculiar velocity surveys, both individually and combined. The analysis is used to optimise the survey design and to predict the accuracy with which cosmological parameters can be constrained. The methods are applied to the 6dFGS but may generally be applied to all redshift and velocity surveys. These analytical results are intended for use without the need for detailed simulations of the surveys involved. It is however recognised that an alternative to this analytical approach is to calculate expected errors directly from Monte Carlo simulations.

The chapter is laid out as follows. In Section 3.3 the Fisher matrix is discussed and written in terms of power spectra. Section 3.4 describes the galaxy and velocity fields and how they relate to the underlying density field. This requires a model for galaxy biassing and redshift space distortions. In Section 3.5 a set of cosmological parameters is chosen as a fiducial model. Section 3.6 discusses the redshift survey on its own. The survey design is modelled and then optimised. The optimised design is then used to predict the best possible constraints on various sets of cosmological parameters. A similar analysis is performed for the velocity survey in

Section 3.7. Finally a combination of the surveys is considered. The Fisher matrix for this combined data set is constructed in Section 3.8 and the results obtained from it are presented in Section 3.9.

3.2 The Gaussian assumption for the density field

This Chapter uses a common assumption that the density field of perturbations follows a Gaussian distribution. There are however, caveats to this approach which should be discussed in order to understand the limitations of this approximation. The primordial density field is expected to have an underlying Gaussian probability distribution since in the inflationary scenario, different Fourier modes enter the horizon with random phases and the resulting superposition is therefore the sum of many independent variates. The central limit theorem then implies that such a sum will have a Gaussian distribution regardless of the underlying probability distributions of the individual modes. At later times however, this Gaussian distribution becomes skewed by gravitational instability. Overdensities become greater at the expense of underdensities. Thus in the present epoch the probability distribution of the density field is only approximately Gaussian and in fact some information must be present in the hierarchy of higher–order moments of the distribution.

As an example of an alternative probability distribution for the density field, consider the log normal distribution. One obvious problem with the Gaussian assumption, is that it assigns a non–zero probability to densities less than zero. One way of enforcing a positive density distribution, as described by Coles and Jones (1991), is to assume a log–normal distribution for the density, such that:

$$\rho(\mathbf{r}) = \exp[X(\mathbf{r})], \tag{90}$$

where the exponent $X(\mathbf{r})$ is a Gaussian variate. There are many motivations for expressing the density in this form – for instance if the velocity fluctuations had a Gaussian distribution, then a log–normal distribution for ρ would follow automatically from the continuity equation.

Although many possibilities exist for the probability distribution of the density field, the Gaussian distribution has the advantage of being relatively simple, approximately correct, and an analysis of the two point correlation function or power spectrum is statistically complete

for this distribution. With the above reservations duly noted, the remainder of this Chapter assumes the density to follow a Gaussian random field.

3.3 The Fisher Information Matrix

The Fisher information matrix can be used to assess the possibilities of parameter estimation from data sets. It is discussed briefly below and in more depth in Tegmark, Taylor and Heavens (TTH, 1997). Appendix A.2 reviews a useful result from this paper for the calculation of the Fisher matrix.

3.3.1 Statistical theory

If a data set consists of n real numbers $x_1, x_2 ..... x_n$ then it is convenient to define a vector $\mathbf{x}$ in an n–dimensional space whose components are x_i. Let $L(\mathbf{x}|\boldsymbol{\theta})$ be the probability distribution of obtaining a data vector $\mathbf{x}$, which also depends on a vector of m model parameters, $\boldsymbol{\theta}$, such that

$$\boldsymbol{\theta} = (\theta_1, \theta_2, ... \theta_m). \tag{91}$$

If $\mathbf{x}$ is a specific data set being considered and $\boldsymbol{\theta}$ is allowed to vary, then the function $L(\mathbf{x}|\boldsymbol{\theta})$ is the Likelihood. Assume the Likelihood is distributed as a Gaussian

$$L(\mathbf{x}|\boldsymbol{\theta}) = \frac{\exp(-\frac{1}{2}\mathbf{x}\mathbf{C}^{-1}\mathbf{x}^t)}{(2\pi)^{N/2}\sqrt{\det \mathbf{C}}} \tag{92}$$

where $\mathbf{C}$ is the data covariance matrix. This expression is true for data with zero mean. If $\boldsymbol{\theta}_0$ represents the *true* parameter values then a good estimate of $\boldsymbol{\theta}$ should be unbiased:

$$\langle \boldsymbol{\theta} \rangle = \boldsymbol{\theta}_0 \tag{93}$$

after many estimates, and should minimize the standard deviations for each value:

$$\triangle\theta_i = (\langle \theta_i^2 \rangle - \langle \theta_i \rangle^2)^{1/2}. \tag{94}$$

The Fisher information matrix is defined as the curvature of the log likelihood about its maximum in parameter space. If all model parameters have an equal prior probability, and if all data vectors have equal prior probability, then Bayes' theorem states that $L(\mathbf{x}|\boldsymbol{\theta}) = L(\boldsymbol{\theta}|\mathbf{x})$. It

is then possible to perform a Taylor expansion of the log likelihood to second order about its maximum in parameter space:

$$\ln L(\boldsymbol{\theta}|\mathbf{x}) \approx \ln L(\boldsymbol{\theta}_0|\mathbf{x}) + \sum_i \Delta\theta_i \frac{\partial}{\partial\theta_i} \ln L(\boldsymbol{\theta}_0|\mathbf{x}) + \sum_{ij} \frac{1}{2}\Delta\theta_i\Delta\theta_j \frac{\partial^2}{\partial\theta_i\partial\theta_j} \ln L(\boldsymbol{\theta}_0|\mathbf{x}). \quad (95)$$

By definition of the maximum likelihood point, the likelihood gradient evaluated at $\boldsymbol{\theta}_0$ is zero and so the second term vanishes. The last term is the second derivative of the log-likelihood evaluated at the maximum likelihood point, which is equivalent to its expectation value – i.e. ensemble averaged. Taking the exponential of both sides of equation (95) shows that the likelihood takes the form of a multivariate Gaussian distribution, and that the inverse of the parameter covariance matrix is given by the Fisher information matrix defined as:

$$\mathbf{F}_{ij} = \left\langle \frac{\partial^2(-\ln L)}{\partial\theta_i\partial\theta_j} \right\rangle. \quad (96)$$

The ensemble average performed here is the equivalent of taking the average of many independent Monte-Carlo simulations of the data. Equation (95) therefore implies that:

$$\langle \Delta\theta_i \Delta\theta_j \rangle = \mathbf{F}_{ij}^{-1}. \quad (97)$$

This means the correlation coefficient between two parameters (labelled i and j) can be written

$$\gamma_{ij} \equiv \frac{\langle \theta_i\theta_j \rangle}{\sqrt{\langle\theta_i^2\rangle\langle\theta_j^2\rangle}} = \frac{\mathbf{F}_{ij}^{-1}}{\sqrt{\mathbf{F}_{ii}^{-1}\mathbf{F}_{jj}^{-1}}}. \quad (98)$$

The Fisher matrix can also be used to place a lower bound on the accuracies of parameter estimates. The Cramér–Rao inequality states that the minimum standard deviation on the value of model parameter i, is

$$\triangle\theta_i \geq \frac{1}{\sqrt{\mathbf{F}_{ii}}} \quad (99)$$

if all other parameters are known, and

$$\triangle\theta_i \geq \sqrt{\mathbf{F}_{ii}^{-1}} \quad (100)$$

if they are unknown (Kendall and Stuart (1969)). This is derived in Appendix A.1. This is the highest possible accuracy to which parameters may be quoted from the data. If an ideal maximum likelihood analysis is performed then the inequality in equation (100) becomes an equality. The Cramér–Rao inequality is also useful as a check on error estimates. If errors are estimated that are smaller than this value then something is wrong.

3.3.2 The Fisher matrix in terms of power spectra

The data vector used here for the redshift survey is the overdensity field expressed in Fourier modes $\mathbf{x} = (\delta(\mathbf{k}_1), \delta(\mathbf{k}_2), ...\delta(\mathbf{k}_m))$. Similarly for the velocity survey the data set is the radial gradient of the radial peculiar velocity in Fourier space. Both the velocity and the density field are best expressed as continuous rather than discrete data and so the Fisher matrix may be written

$$\mathbf{F}_{ij} = \frac{1}{2}\int d^3r \frac{d^3k}{(2\pi)^3}\frac{\partial}{\partial\theta_i}\ln C(k)\frac{\partial}{\partial\theta_j}\ln C(k) \tag{101}$$

(Taylor and Watts (2001)) where $C(k)$ is the covariance of modes of wavenumber k. This result is true if the underlying fields – the overdensity and velocity distributions – are Gaussian random fields. In the real world of course, the data will be subject to noise. It is through modelling this noise as a function of radius that key survey parameters may be optimised. It is convenient to place this noise term directly into the model power spectrum: $C(k) \approx P(k) + N(r)$ so

$$\mathbf{F}_{ij} = \frac{1}{2}\int \frac{d^3k}{(2\pi)^3}\partial_i \ln P(\mathbf{k})\partial_j \ln P(\mathbf{k}) V_{\rm eff}(\mathbf{k}), \tag{102}$$

where

$$V_{\rm eff}(\mathbf{k}) = \int d^3r \left(\frac{P(\mathbf{k})}{P(\mathbf{k}) + N(\mathbf{r})}\right)^2 \tag{103}$$

can be thought of as the effective volume of the survey. The above analysis applies generally to the galaxy power spectrum, the velocity power spectum and the covariance of a combination of the two data sets. In the following analysis, we consider only the linear galaxy power spectrum as we wish to combine it with the linear velocity power spectrum. We also truncate the analysis at a value $k_{max} \approx 0.2h{\rm Mpc}^{-1}$ which is roughly the limit of the linear regime. This is appropriate for predicting the results of an analysis which only considers a data vector with wavenumbers below this value.

3.4 Galaxy density and radial velocity fields

We shall assume for our analysis that the relevant field is the linear galaxy redshift-space density perturbation in the plane–parallel approximation. We write

$$\delta_g^s(\mathbf{k}) = D(k\sigma_v\mu)[(b_L + \mu^2 f)\delta_m(\mathbf{k}) + \epsilon] \tag{104}$$

where f is the growth index of density perturbations, b_L is a linear bias parameter, $\delta_m(\mathbf{k})$ is the linear matter density field, $\mu = \hat{\mathbf{r}}.\hat{\mathbf{k}}$ is the cosine angle between the wave–vectors of the density field and the observers line of sight and ϵ is a stochastic bias term. The first term in the square brackets is the linear redshift distortion in the plane parallel approximation. This is discussed in Section 1.11.6 and results in the linear distortion effect described by equation (76). The growth index of perturbations f, is discussed in Section 1.10.3 and is generally dependent on both the matter density parameter and the cosmological constant. Here we use the approximation of Peebles (1980): $f \equiv d\ln\delta/d\ln a \approx \Omega_m^{0.6}$. We have also included a nonlinear redshift distortion term

$$D(k\sigma_v\mu) = (1 + k^2\sigma_v^2\mu^2/2)^{-1/2} \tag{105}$$

where σ_v is the pairwise radial velocity dispersion. This accounts for the Fingers of God effect discussed in Section 1.11.6. The form of the function is motivated by observations of the redshift space correlation function. The stochastic bias term, ϵ in equation (104), is a random variable which accounts for all of the physics of galaxy biassing which are not adequately described by a simple linear biassing relation (Dekel and Lahav (1999)). We assume $\langle\epsilon\rangle = 0$ and $\langle\epsilon^2\rangle = \sigma_\epsilon^2$. We define a second bias parameter as the ratio of power in the galaxy distribution $P_{gg}(k)$, to that in the matter fields, $P_{mm}(k)$:

$$b^2 = \frac{P_{gg}(k)}{P_{mm}(k)} = b_L^2 + \frac{\sigma_\epsilon^2}{P_{mm}(k)}. \tag{106}$$

We may also introduce a galaxy correlation coefficient, r_g, defined by

$$r_g = \frac{P_{gm}(k)}{\sqrt{P_{gg}(k)P_{mm}(k)}} = \frac{b_L}{b} \tag{107}$$

(Dekel and Lahav (1999)) which will allow us to change between the two bias parameters. Although b_L is more physical, it is b which is more commonly measured.

We choose our second field to be the radial gradient of the radial peculiar velocity field, defined in Fourier space as

$$u'(\mathbf{k}) = \frac{\partial}{\partial r}\hat{\mathbf{r}}.\mathbf{v}(\mathbf{k}) = -\mu^2 H f(\Omega_m)\delta_m(\mathbf{k}) \tag{108}$$

where H is the Hubble parameter. Again we assume that the velocity field can be described as a plane wave, invoking the distant observer approximation. The modes of the radial velocity are uncorrelated with the modes of the density field. However the Fourier space radial gradient of the radial velocity *is* correlated and will yield a cross power spectrum.

The auto– and cross–power spectra of these two fields are

$$P^s_{gg}(\mathbf{k}) = D^2(k\sigma_v\mu)(1 + 2\mu^2 r_g\beta + \mu^4\beta^2)b^2 P_{mm}(k), \quad (109)$$

$$P_{u'u'}(\mathbf{k}) = \mu^4 H^2\beta^2 b^2 P_{mm}(k), \quad (110)$$

$$P^s_{gu'}(\mathbf{k}) = -\mu^2 H\beta D(k\sigma_v\mu)(r_g + \mu^2\beta)b^2 P_{mm}(k), \quad (111)$$

where

$$\beta \equiv \frac{f(\Omega_m)}{b} \quad (112)$$

is the linear redshift distortion parameter. These are the three spectra to be measured: the redshift–space galaxy power spectrum, $P^s_{gg}(\mathbf{k})$, the radial velocity gradient power spectra, $P_{u'u'}(\mathbf{k})$, and the cross–spectra of these fields, $P^s_{gu'}(\mathbf{k})$.

The noise terms associated with these fields are

$$N_{gg}(r) = \frac{1}{n_g(r)}, \quad (113)$$

$$N_{u'u'}(r) = \mu^2 k^2 \frac{\sigma^2_{DI}(r)}{n_v(r)}, \quad (114)$$

$$N_{gu'}(r) = 0, \quad (115)$$

where $n_g(r)$ and $n_v(r)$ are the number densities of the galaxy and velocity surveys, respectively. The factors of μ^2 and k^2 are due to our use of the radial velocity gradient. We quantify the number densities in Sections 3.6 and 3.7.1. The noise term for radial velocities arises from the intrinsic uncertainty on galaxy positions due to the dispersion in the D_n–σ relation for ellipticals. This can be approximated by

$$\sigma_{DI}(r) = \sigma_0 H r e^{r/r_{\rm err}}, \quad (116)$$

where σ_0 is the fractional error on the distance indicator. The exponential factor in the above equation is designed to model the way that the D_n–σ errors will blow up beyond some limiting depth. This modification to more accurately describe the real D_n–σ errors was requested by the 6dFGS team (Will Saunders, private communication). We allow the fractional distance error to blow up beyond some fiducial distance, $r_{\rm err}$. We choose a conservative $r_{\rm err}$ of $135\,h^{-1}\,{\rm Mpc}$ (which is 90 % of the 6dF median depth), but allowing this to decrease to as low as 50% of the median depth had little effect on our results.

3.5 The matter power spectrum and cosmological parameters

Using the Fisher matrix and the Cramér–Rao inequality we can estimate how successfully a maximum likelihood analysis will constrain model parameters in the power spectrum. We define the matter power spectrum to be

$$P_{mm}(k) = Q^2 k^n T^2(k; \Gamma, \omega_b) \tag{117}$$

where Q is the amplitude, n is the spectrum of the primordial power spectrum and T is the transfer function. We use the transfer function given by Eisenstein and Hu (1998). This depends on the shape parameter

$$\Gamma = \Omega_m h \tag{118}$$

which stretches the scale of the transfer function, and

$$\omega_b = \Omega_b h^2 \tag{119}$$

where Ω_b is the density parameter of baryonic matter. In terms of Q the amplitude of galaxy clustering and the amplitude of the velocity field are respectively:

$$A_g = bQ \tag{120}$$

$$A_v = f(\Omega_m)Q \tag{121}$$

where b is the simple biassing parameter and $f(\Omega_m)$ is the growth rate of perturbations from Section 1.10.3. Finally we can make use of the redshift space distortions (Section 1.11.6) to constrain the parameter $\beta = \Omega^{0.6}/b$ and the mass–galaxy correlation coefficient r_g. We assume a scale invariant primordial power spectrum $n = 1$ and leave $(A_g, A_v, \Gamma, \omega_b, b)$ as potential free parameters in a future likelihood analysis . We assume as fiducial values for these parameters $(A_g, A_v, \Gamma, \omega_b, \beta) = (5 \times 10^{-5}, 0.486 \times 5 \times 10^{-5}, 0.195, 0.025, 0.486)$.

3.6 The Galaxy Redshift Survey

3.6.1 Survey dependence

In the Fisher matrix analysis the details of the survey are incorporated in the noise term in equation (102). Here we assume the noise to be shot noise caused by the discrete nature of

the data. To model the shot noise as a function of the survey depth we use a parameterised expression for the mean radial density of galaxies in the sample:

$$n_g(r) = N_g \frac{a}{r_*^3 \Gamma[3/a]} e^{-(r/r_*)^a} \tag{122}$$

where r_* is the effective depth of the survey, N_g is the total number of galaxies and typically $a = 1.5$. The mean depth of the survey is $r_{\rm m} = 1.5 r_*$ and the number of galaxies roughly scales as

$$N_g \approx 4 \times 10^{-2} \alpha r_{\rm m}^{3.125}, \tag{123}$$

from fitting the near infrared differential number counts of Kochanek et al. (2001). To simplify the calculation we have approximated their median redshifts as mean redshifts. The sampling fraction, α, is simply the fraction of sources expected at any redshift which are actually observed. The three survey parameters are then $f_{\rm sky}$, r_* and α.

We approximate space to be Euclidean. This assumption is necessary to allow the more complicated supermatrix programs, described in the later sections, to run in reasonable times. At the median redshift of the survey, $z = 0.055$, the proper distance for our fiducial cosmology is $r = 163 h^{-1}{\rm Mpc}$ whereas the Euclidean distance gives $r = 165 h^{-1}{\rm Mpc}$. At twice the median redshift, the proper distance is $r = 322 h^{-1}{\rm Mpc}$ whereas the Euclidean distance is $r = 330 h^{-1}{\rm Mpc}$, corresponding to $\sim 2\%$ difference.

3.6.2 Survey optimisation

Heavens and Taylor (1997) show that optimisation of a galaxy redshift survey for a fixed timescale reduces to a single parameter problem. We choose to optimise the survey for the parameter A_g. The minimum variance bound on an estimate of the single parameter A_g is given by the Cramér–Rao inequality (equation (100)) as

$$\frac{\Delta A_g}{A_g} = \left[\frac{1}{2} \int \frac{d^3k}{(2\pi)^3} V_{\rm eff}^g(k,\mu) \right]^{-1/2}. \tag{124}$$

We assume that the survey timescale will depend on the survey parameters according to

$$t = t_0 \alpha f_{\rm sky} (r_m / h^{-1}{\rm Mpc})^7. \tag{125}$$

For the 6dFGS the ratio t/t_0 is calculated as 7×10^{17} assuming the expected sky fraction and depth. With the time constraint the survey depth scales as

$$r_m = 150(\alpha f_{\rm sky})^{-1/7} h^{-1}{\rm Mpc}. \tag{126}$$

As we increase the sampling fraction the achievable depth for a fixed timescale survey decreases. Figure 9 shows the fractional error bars on A_g calculated from the Fisher matrix for a range of depths and sampling fractions. The four lines correspond to fixed fractions of sky coverage – the lowest line is for an all sky survey and it is clear that for all combinations of r_* and α the largest solid angle possible should be chosen. Having fixed this sky fraction – 6dFGS should lie somewhere above the third highest line in Figure 9 – we can look for the minimum point to find the optimal combination of r_m and α. For a survey covering a hemisphere we find a median depth of $r_m = 230h^{-1}$Mpc and $\alpha \approx 0.16$ to be optimal although we find the curves to be rather flat around the optimal sampling fraction. The figure shows the 6dFGS design as a red dot and the optimal design as a black dot. The flatness of the line indicates that the 6dFGS design is satisfactory in terms of parameter estimation and it may be that other factors are more important – such as keeping the survey depth low so as to increase the availability of good spectra for the velocity survey.

3.6.3 Power, Noise and Volume

Figure 10 shows our model power spectrum assuming the shot noise of the 6dFGS and assuming optimal design. The shaded region corresponds to the uncertainty on the amplitude of the power spectrum in wavenumber bins of width $\Delta \ln k = 0.5$. This is calculated analytically from the Fisher matrix as (Feldman et al. (1994)):

$$\overline{\Delta P}(k) = \frac{2\pi P(k)}{\sqrt{k^3 \Delta \ln k V_{\rm eff}(k)}}. \tag{127}$$

Also plotted in Figure 10 is the effective volume of the survey. The peak of this curve indicates the scale at which the information content is maximised. The limit of linear theory is roughly $k \sim 0.2$ so the information content maximizes just into the non–linear regime.

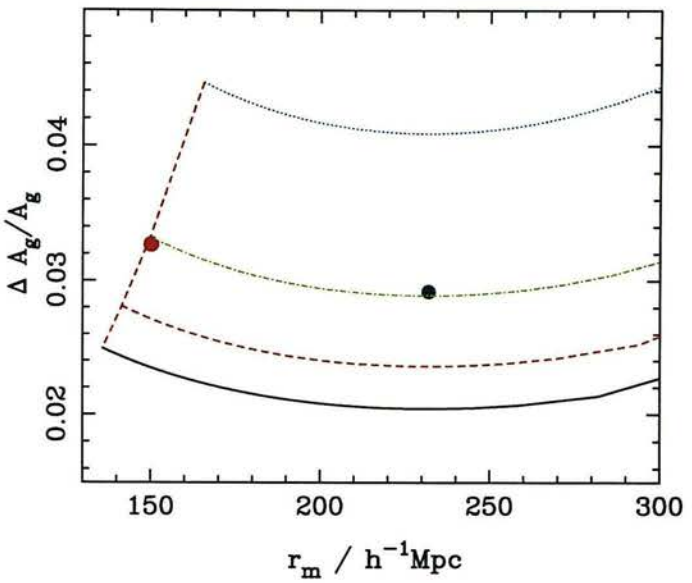

Figure 9: *Optimising the redshift survey. The four lines correspond to different fractions of sky covered: a quarter, a half, three quarters and all–sky from top to bottom. The red circle shows the 6dFGS survey design whereas the black circle is the optimal design. Although the 6dFGS design is not optimal the flatness of the line indicates that this will not have a big effect on parameter estimation.*

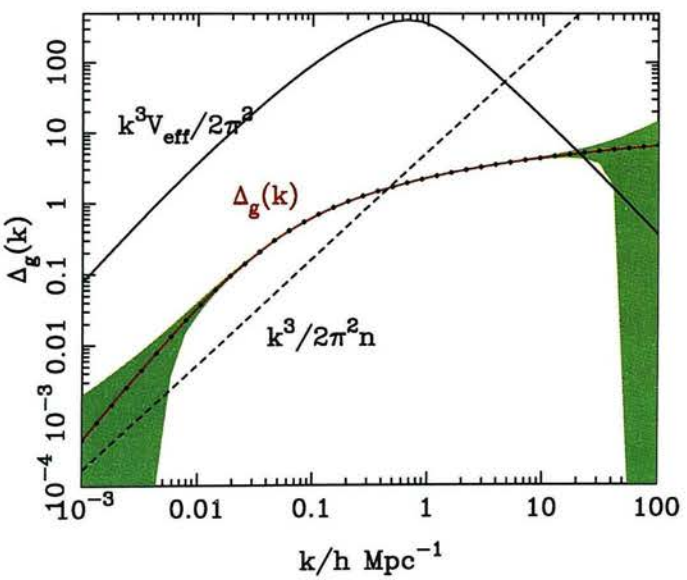

Figure 10: *Power, effective volume and noise, calculated for the 6dF redshift survey. The survey is parameterised by its sky coverage, depth and total number of galaxies.*

3.6.4 Derivatives of the matter power spectrum

The obvious next stage is to determine the information content of the power spectrum for the different parameters discussed in Section 3.5. From the form of the Fisher matrix (equation 102) it can be seen that the derivatives of the power spectrum with respect to these parameters quantify the information about them in $P(k)$. Figure 11 shows the derivatives as a function of wavenumber, again for the 6dF redshift survey. This tells us relatively how accurately each parameter could be individually constrained. The shapes of the derivatives are also interesting. If these functions are visualised as vectors in a Hilbert space, then the Fisher matrix takes the scalar product of the vectors. The more similar the shapes of the lines, the more parallel the functions are – i.e. the parameters are more strongly correlated. In the extreme of two parameters being completely degenerate, the curves are identical, the Fisher matrix becomes singular and the resulting individual uncertainties become infinite. From Figure 11 we can see that A_g, β and r_g all have very similar shapes. All three parameters appear in the normalisation of the power spectrum leading to a strong correlation between them. The remaining two parameters, Γ and ω_b, also have similar derivatives. The shape Γ parameterises the break scale between modes suppressed during radiation domination and unsuppressed modes. The baryon fraction ω_b characterises the damping effect of baryons on the high k end of the spectrum. The two parameters describe different physical effects which produce a similar damping of the spectrum. Peacock and Dodds (1994) show that the effect of baryons can be described by an effective shape parameter:

$$\Gamma_{\rm eff} = \Gamma e^{-2\Omega_b h}, \tag{128}$$

which models both effects and explicity shows the degeneracy between the two parameters. Figure 11 shows that the matter power spectrum contains more information on the shape than the baryon content.

3.6.5 Analytical results for A_g and β

In the limit of negligible shot noise and assuming all other parameters are known, we can calculate analytical expressions from the two–by–two Fisher matrix for A_g and β for any general

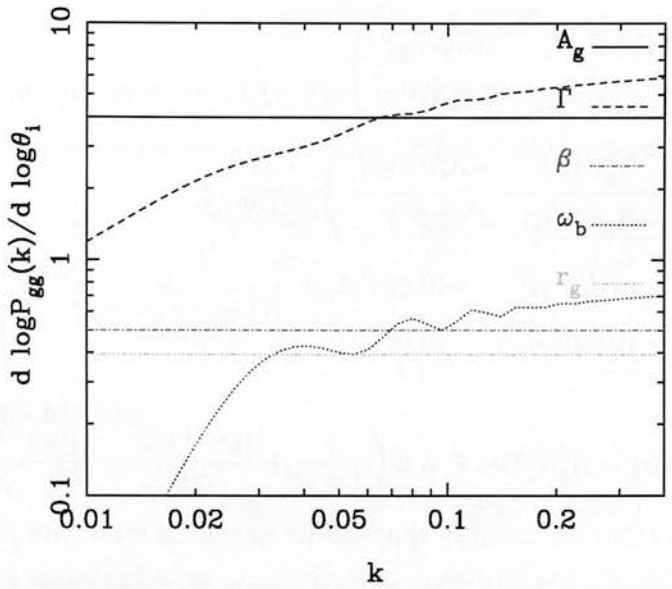

Figure 11: *Derivatives of the matter power spectrum, for the 6dF redshift survey.*

redshift survey. If we assume a constant effective volume V, the Fisher matrix becomes,

$$\mathbf{F}_{ij} = \frac{1}{2} V \int \frac{d^3k}{(2\pi)^3} \, \partial_i \ln P^s_{gg}(k,\mu) \, \partial_j \ln P^s_{gg}(k,\mu). \tag{129}$$

If galaxies and matter are completely correlated, $r_g = 1$ and the redshifted power spectrum becomes:

$$P^s_{gg}(k) = D^2(k\sigma\mu)(1 + 2\mu^2\beta + \mu^4\beta^2) b^2 P_{mm}(k). \tag{130}$$

The Fisher matrix can be found analytically by differentiating this and averaging over angles:

$$\begin{aligned} \mathbf{F} &= \begin{pmatrix} \mathbf{F}_{A_g A_g} & \mathbf{F}_{\beta A_g} \\ \mathbf{F}_{\beta A_g} & \mathbf{F}_{\beta\beta} \end{pmatrix} \\ &= \begin{pmatrix} 2/A_g^2 & \Phi(\beta)/A_g\beta \\ \Phi(\beta)/A_g\beta & \Theta(\beta)/\beta^2 \end{pmatrix} \frac{k^3 V}{6\pi^2} \end{aligned} \tag{131}$$

where

$$\Theta(\beta) = \int_{-1}^{+1} d\mu \left(\frac{\beta\mu^2}{1+\beta\mu^2} \right)^2 = 2 + \frac{1}{1+\beta} - 3\frac{\tan^{-1}\sqrt{\beta}}{\sqrt{\beta}}, \tag{132}$$

$$\Phi(\beta) = \int_{-1}^{+1} d\mu \left(\frac{\beta\mu^2}{1+\beta\mu^2} \right) = 2\left(1 - \frac{\tan^{-1}\sqrt{\beta}}{\sqrt{\beta}} \right). \tag{133}$$

The parameter covariance matrix is the inverse of the Fisher matrix:

$$\begin{aligned}\mathbf{F}^{-1} &= \begin{pmatrix} \langle(\Delta A_g)^2\rangle & \langle\Delta\beta\Delta A_g\rangle \\ \langle\Delta\beta\Delta A_g\rangle & \langle(\Delta\beta)^2\rangle \end{pmatrix} \\ &= \begin{pmatrix} A_g^2\Theta(\beta) & -A_g\beta\Phi(\beta) \\ -A_g\beta\Phi(\beta) & 2\beta^2 \end{pmatrix} \frac{6\pi^2}{k^3 V\Psi(\beta)} \\ &= \begin{pmatrix} (0.01)^2 A_g^2 & -(0.02)^2 A_g\beta \\ -(0.02)^2 A_g\beta & (0.06)^2\beta^2 \end{pmatrix} \left(\frac{k}{0.2h\mathrm{Mpc}^{-1}}\right)^{-3} \left(\frac{V}{10^8[\,h^{-1}\mathrm{Mpc}]^3}\right)^{-1} \end{aligned} \quad (134)$$

where

$$\Psi(\beta) = A_g^2\beta^2\,\mathrm{Det}\,\mathbf{F} = 2\left(\frac{1}{1+\beta} + \frac{\tan^{-1}\sqrt{\beta}}{\sqrt{\beta}} - 2\frac{[\tan^{-1}\sqrt{\beta}]^2}{\beta}\right) \quad (135)$$

and in equation (134) we calculate approximate numerical results for $\beta = 0.5$. For a survey with a volume of order $\sim 10^8[h^{-1}\mathrm{Mpc}]^3$, in the absence of shot noise, we obtain marginalised fractional error bars of $\Delta A_g/A_g = 0.01$ and $\Delta\beta/\beta = 0.057$ at a wavenumber of $k \approx 0.2$. When β is estimated on its own the uncertainty reduces to,

$$\frac{\Delta\beta}{\beta} = \left(\frac{V}{4.5\times 10^4[h^{-1}\mathrm{Mpc}]^3}\right)^{-1/2}. \quad (136)$$

It must be remembered that this is in the unrealistic case of negligible shot noise. This result says that even for a very well sampled survey, the effective volume has to be large to get a tight constraint on β. For instance to achieve a 1% error requires a survey volume of $V \approx 4.5\times 10^8[h^{-1}\mathrm{Mpc}]^3$ over four times larger than either the 2DF, 6DF or SDSS redshift surveys. For these surveys we can expect an accuracy of $\Delta\beta/\beta \approx 0.02$. One alternative to increasing the volume is to move to the non–linear regime as outlined by Taylor and Watts (2001). Another possibility as discussed later in this chapter is to combine data sets.

The correlation between the parameters may also be derived from equation (134) in the limit of negligible shot noise. For $\beta = 0.5$ we find $\gamma_{A_g\beta} = -0.78$ which is consistent with the value found by Taylor et al. (2001b) for the PSCz.

3.6.6 The amplitude of mass clustering

Taylor et al. (2001b) and Tadros et al. (1999) suggest that in the absence of stochastic biassing between mass and galaxies the amplitude of mass clustering can be calculated from the com-

bination:

$$Q = (A_g\beta)\Omega_m^{-0.6} \propto \sigma_8\Omega_m^{-0.6} \tag{137}$$

which can be estimated more accurately than any of its constituents because it marginalises over the longest axis of the parameter covariance ellipse. From propagation of errors we find:

$$\begin{aligned}\frac{\Delta Q}{Q} &= \sqrt{\frac{\langle(\Delta A_g)^2\rangle}{A_g^2} + \frac{\langle(\Delta\beta)^2\rangle}{\beta^2} + 2\frac{\langle\Delta A_g\Delta\beta\rangle}{A_g\beta}} \\ &= \sqrt{\frac{6\pi^2}{k^3V}}\left(\frac{\Theta(\beta)+2-2\Phi(\beta)}{\Psi(\beta)}\right).\end{aligned} \tag{138}$$

When $\beta = 0.5$ this becomes

$$\frac{\Delta Q}{Q} = 0.05\left(\frac{k}{0.2h^{-1}\mathrm{Mpc}}\right)^{-3}\left(\frac{V}{10^8[h^{-1}\mathrm{Mpc}]^3}\right)^{-1}. \tag{139}$$

The marginalised uncertainties on A_g and β are 1% and 6% and the uncertainty on Q is 5%. This is a better constraint than β but not A_g. Because the error bars in A_g and β are so different, the major axis of the error ellipse does not lie at 45 degrees to the parameter axis and so the marginalisation does not allow Q to be better constrained than both of the independent parameters.

3.6.7 Three parameter set: A_g, Γ and β

The more free parameters we have the more complicated the correlations between them become and uncertainties increase. To start off with, we look at the three parameters A_g, Γ and β. Figure 12 (upper panel) shows the expected marginalised errors from a joint constraint of the three parameters, from the 6dF redshift survey. The errors represent the minimum variance bound calculated from the Fisher matrix, and are plotted as a function of the maximum wavenumber analysed. The linear theory we use here is robust to a wavenumber of $k \sim 0.2\,h^{-1}\mathrm{Mpc}$ so it is from here that we quote error bars. The best estimated parameter is Γ which is constrained to within about 2% from the 6dF redshift survey. The uncertainty on A_g and β is about 3%, which is close to the conditional estimate predicted in Section 3.6.5. The lower panel of Figure 12 shows the correlation between the parameters, obtained from the Fisher matrix using equation (98). A_g and Γ are strongly anti–correlated. This is because a change in amplitude can be mimiced by a shift in the break scale. In Section 3.6.5 it was shown that for a two parameter fit A_g and β are strongly anti–correlated. From Figure 12 we can see that this correlation has

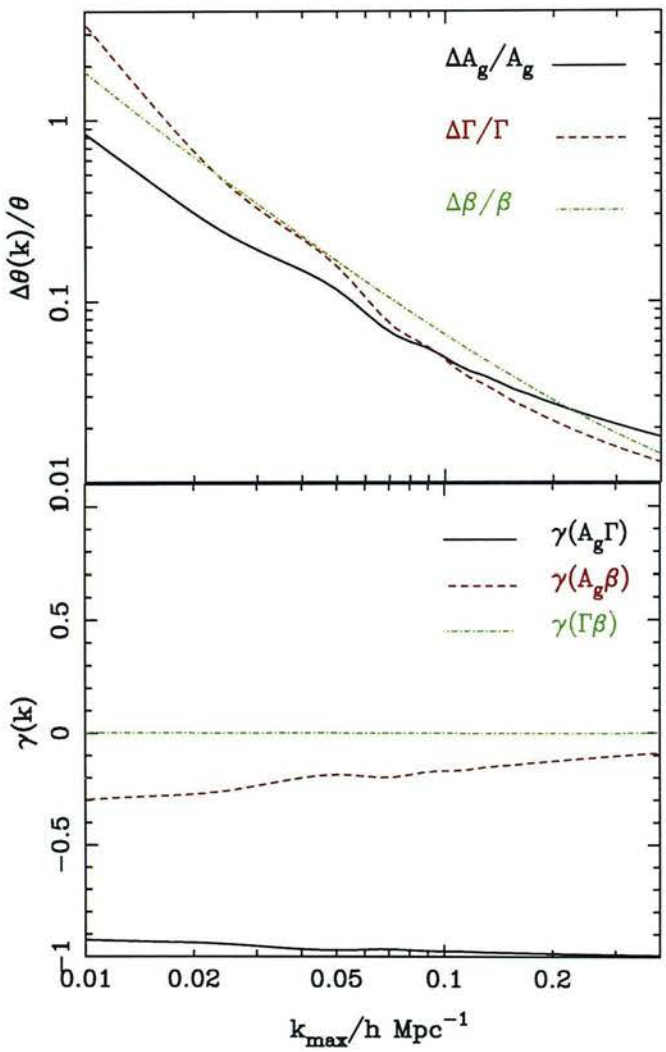

Figure 12: **Upper plot:** *The expected fractional uncertainty on A_g, Γ and β from the 6dF redshift survey.* **Lower plot:** *The correlations between A_g, Γ and β*

weakened with the introduction of a third parameter. A final comment about Figure 12 is that using the Fisher matrix values for β and A_g it is possible, as in Section 3.6.6, to calculate the expected uncertainties on Q, the amplitude of mass clustering. This time, including shot noise and a more accurate value for the 6dFGS effective volume we find $\frac{\Delta Q}{Q} = 4.8\%$.

3.6.8 Four parameter set: A_g, Γ, β and ω_b

Figure 13 shows an equivalent analysis for a four parameter set involving ω_b. The introduction of ω_b has little effect on the uncertainties on A_g and β for wavenumbers greater than $k > 0.05h\mathrm{Mpc}^{-1}$. The top half of the lower panel in Figure 13 shows that this is because the correlations between ω_b and these parameters is small. However there *is* a strong correlation between ω_b and Γ, $\gamma(\Gamma\omega_b) \approx 0.9$, which is expected from the discussion in Section 3.6.4. Because of this correlation, the presence of ω_b as a free parameter degrades the uncertainty on Γ. In any case the uncertainty on ω_b itself is dissapointingly large, $\Delta\omega_b/\omega_b \approx 0.45$, reflecting its weak effect on the matter power spectrum. As with all the parameters, the main way to improve on sensitivity is with a larger redshift survey.

One possible improvement would be to use the degenerate parameter

$$f_b = \frac{\Omega_b}{\Omega_m} \propto \frac{\omega_b}{\Gamma}, \tag{140}$$

the baryon fraction. Since ω_b and Γ are correlated, use of this combination would involve marginalising over the longest axis of the ω_b–Γ error ellipse, making the fractional error on f_b smaller than the constraint of either of the parameters independently. Unfortunately, although this is true for a two parameter analysis, in the case of four free parameters, other correlations come into play, complicating matters. When the analysis is repeated using this new parameter there is no significant improvement. The main limiting factor here is the survey volume. In order to reduce the error on ω_b to the 10% level we predict one would require a survey some 10 times larger.

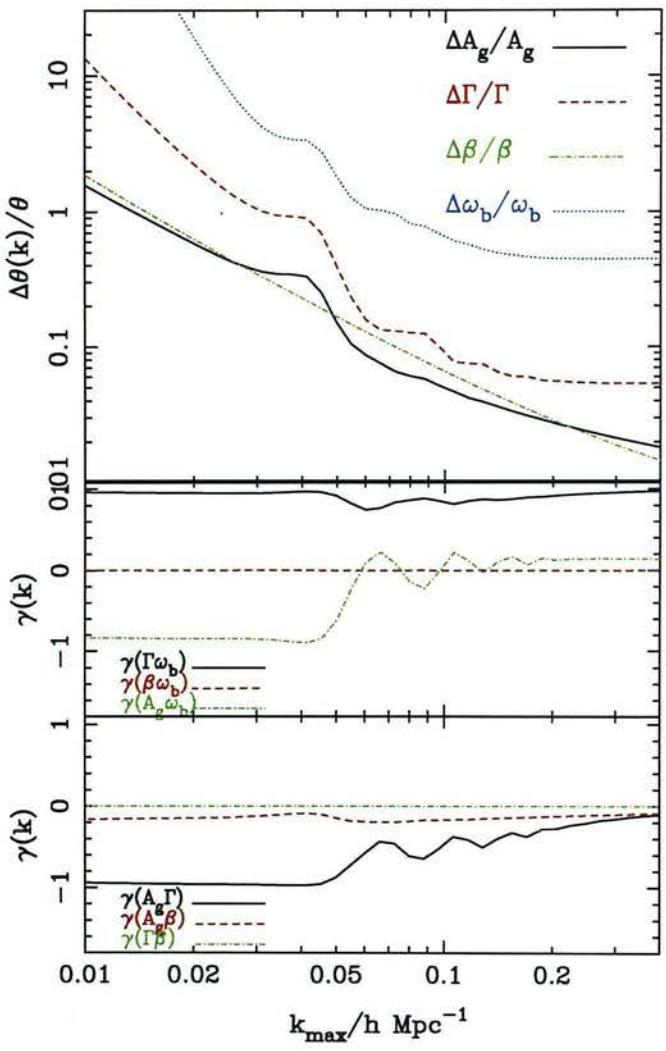

Figure 13: **Upper plot:** *Predicted errors for a four parameter analysis involving ω_b, of the 6dF redshift survey.* **Lower plot:** *Correlations in a four parameter space with ω_b*

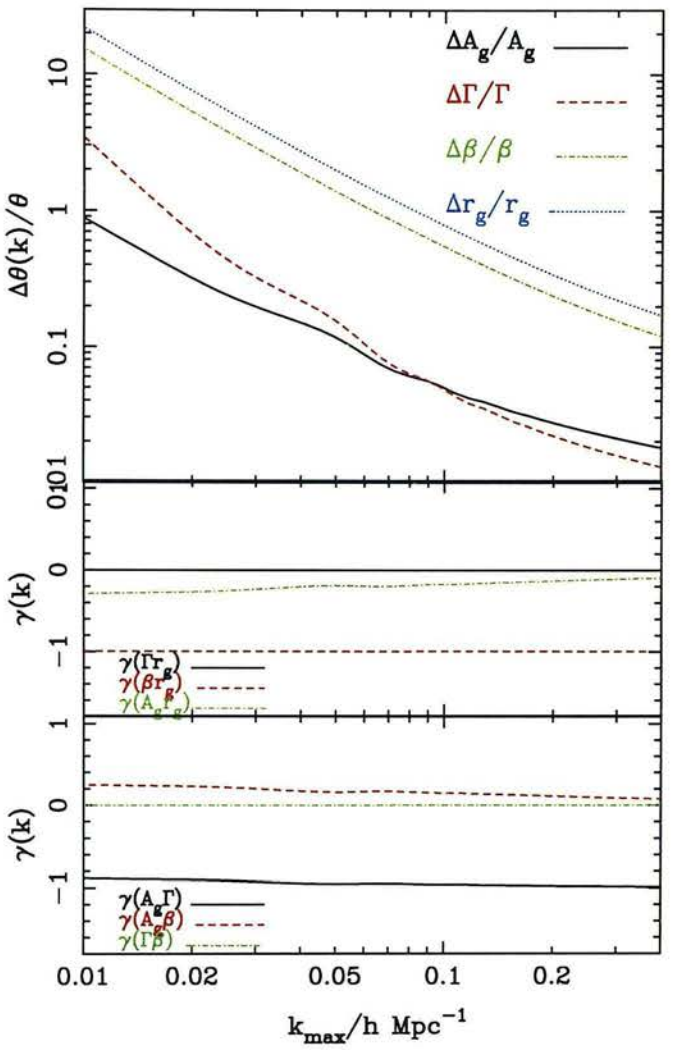

Figure 14: **Lower plot:** *Uncertainties for a four parameter analysis invloving r_g.* **Lower plot:** *The correlations between A_g, Γ, β and r_g*

3.6.9 Four parameter set: A_g, Γ, β and r_g

The next step was to see if r_g would fare any better as a fourth parameter. Figure 14 shows that this time A_g and Γ are unaffected by this new parameter, but the uncertainty on β is dramatically increased, with both the fractional errors on β and r_g around 35%. Again this can be traced to the expected degeneracy between r_g and β discussed in Section 3.6.4, and shown in the lower panel of Figure 14, where $\gamma_{\beta r_g} \approx -1$. Interestingly the correlation between A_g and β has now become positive, so that the error on their combination will be increased.

It seems that from redshifts alone, three parameters can be constrained by the 6dFGS: A_g, Γ, β.

3.7 The Galaxy Velocity Survey

The 6dFGS will produce a relatively large number of peculiar velocities drawn from a uniform selection function over a wide area. It is therefore appropriate to perform a similar Fisher matrix analysis of the velocity power spectrum. This is the first such analysis of a velocity survey and it follows the prescription set out in the previous section: first the survey design is optimised for information content, then predictions are made as to the success of this survey in constraining parameters.

3.7.1 Velocity survey parameters

The velocity survey may be parameterised in a similar way to the redshift survey, using $f_{\rm sky}$ to denote the fraction of sky covered and using a simple empirical model for the mean density of galaxies as a function of radius. The velocity survey will be composed of early type galaxies comprising a subset of the overall galaxy catalogue. The velocity survey is also restricted in depth, cutting off at the mean redshift of the galaxy survey. The radial density distribution of equation (122) is therefore modified in amplitude and restricted in depth, but otherwise maintains the same form. The noise component this time has an extra element. In addition to shot noise, the scatter in the distance indicators follows through to the radial velocities. This scatter is parameterised by σ_0 the expected fractional error in the distance indicator. The parameters describing the velocity survey design are therefore r_m, $f_{\rm sky}$, α_v and σ_0.

Given a model for the noise, an effective volume can be defined in analogy to the effective volume for the density field analysis. Figure (15) shows the 3–D velocity power spectrum for the fiducial LCDM cosmology. Also plotted is the effective volume of the 6dF velocity survey, showing which region of the power spectrum contains the most information.

The power spectrum is that of the three dimensional velocities but the information on the

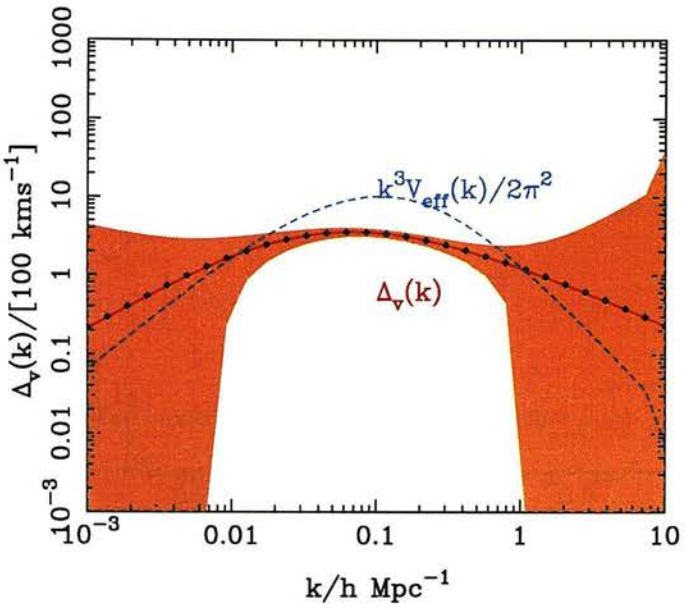

Figure 15: *The peculiar velocity power spectrum,* $\Delta_v(k) = \sqrt{k^3 P_v(k)/2\pi^2}$ *with predicted noise and effective volume for the 6dF peculiar velocity survey. The velocity survey is parameterised by* $f_{\rm sky}$*, cut–off depth,* N_v *and the accuracy of the distance indicator.*

amplitude will come from the data which are only radial velocities. The effective volume plotted here is therefore that of the radial velocity power spectrum. The angular part of this effective volume may be integrated analytically and yields

$$V^u_{\rm eff}(k) = 2\pi f_{\rm sky} \int_0^\infty dr r^2\, \Theta \left[\frac{n_u(r) P_{uu}(k)}{\sigma^2_{DI}(r)} \right] \tag{141}$$

where the function $\Theta(z)$ is defined in equation (132), and σ_{DI} is the absolute error in the distance indicator, given in equation (116). Figure 15 shows that the peak of the effective volume is just within the linear regime. Again this is the point at which information is maximised.

3.7.2 Band Averaged Power

In analogy to the density power spectrum plot (Figure (10)) Figure (15) also shows the error expected on measurements of the band averaged power. The pass bands are chosen to be the same as for the galaxy power spectrum, $\Delta \ln k = 0.5$, and so the uncertainties on the two spectra may be directly compared. The velocity power has far larger uncertainty, mainly because of the smaller size of the data set and hence greater shot noise. The analysis here assumes the optimised survey design, an issue which is discussed next.

3.7.3 Optimisation of a velocity survey

Again optimisation is meant in the sense of maximizing the survey's information about one parameter. Again the uncertainty on the power spectrum amplitude, this time A_v, is minimized for a fixed survey timescale. The fractional error on A_v is simply

$$\frac{\Delta A_v}{A_v} = \left[\int \frac{d\ln k}{4\pi^2} k^3 V_{\rm eff}^v(k) \right]^{-1/2}. \tag{142}$$

The timescale for a velocity survey will be proportional to the total number of sources in the survey, $N \sim r^3$ and the reciprocal of the effective flux of the sources, $S \sim r^{-2}$. It will also scale as the distance error per source, σ_{DI}, the fraction of sky surveyed, $f_{\rm sky}$ and the fraction of sources sampled α_v. So the timescale may be written

$$t \approx \alpha_v f_{\rm sky} \sigma_0^{-2} \left(\frac{r_m}{h^{-1}{\rm Mpc}} \right)^7 t_0 \tag{143}$$

where t_0 is some constant of proportionality. For the proposed 6dF survey plan we find $t/t_0 \approx 2 \times 10^{13}$. The fractional error on A_v can now be minimized in survey–space whilst applying the constant timescale constraint as well as restricting the survey to $f_{\rm sky}$, $\alpha_v \leq 1$. At all depths and for all sampling fractions, the uncertainty on any parameter scales as $\Delta\theta \propto f_{\rm sky}^{-1/2}$, and so all information is scaled according to the sky fraction. It is convenient then to fix $f_{\rm sky}$ and optimise with respect to the remaining survey parameters. Also, the fractional uncertainty in the distance indicator, σ_0, and the source sampling rate, α_v, always appear in the ratio σ_0^2/α_v, further reducing the number of free parameters in the survey design. Finally the depth of the survey can be related to the other parameters by

$$r_m \approx 150 \left[\frac{\sigma_0^2}{\alpha_v f_{\rm sky}} \right]^{1/7} h^{-1}{\rm Mpc}. \tag{144}$$

using equation (143). So the survey can be optimised with respect to the single degenerate parameter

$$\nu = \frac{\sigma_0}{\sqrt{\alpha_v}}. \tag{145}$$

In practice of course, there is a lower limit to the accuracy of the distance indicator. It is therefore more practical to optimise in terms of the parameter

$$\eta = \frac{\sigma_0 + \sigma_{\rm min}}{\sqrt{\alpha_v}} \tag{146}$$

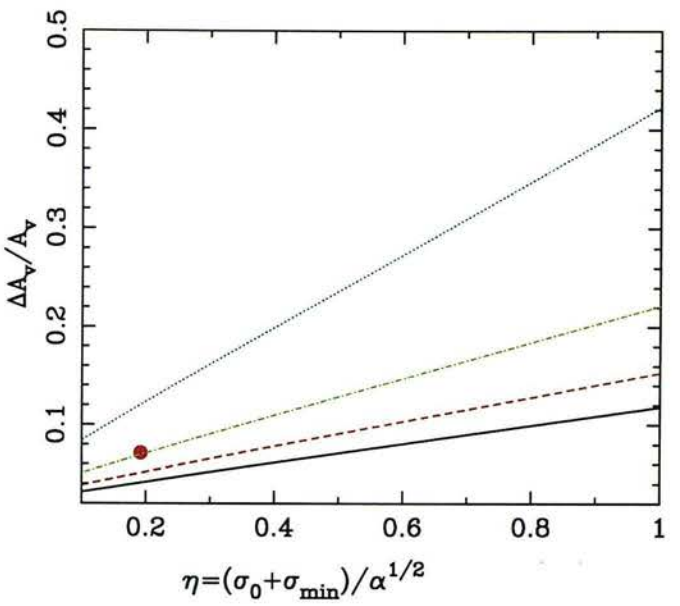

Figure 16: *The normalised fractional uncertainty on the amplitude of the velocity power spectrum for the degenerate parameter* $\eta = \frac{(\sigma_0+\sigma_{\rm min})}{\sqrt{\alpha_v}}$, *for equal–time surveys. The four lines correspond to different sky fractions (see Figure 9). The red dot shows the design of the 6dFGS.*

which better reflects the survey's limitations. Currently the most optimistic estimated accuracies of the D_n–σ relation provide distance estimates with a minimum scatter of $\sigma_{\rm min} \approx 0.1$ e.g. (Jorgensen et al. (1993); Gregg (1995)).

Figure (16) shows the fractional uncertainty on a conditional measurement of A_v for equal–time surveys as a function of η. The different lines correspond to different sky fractions, $f_{\rm sky}$. The lowest, solid line is for an all–sky survey, with higher lines reduced by one quarter. The expected 6dF velocity survey lies somewhere around the red dot. This also reflects the expected precision of its distance indicators.

One other parameter which has not previously been mentioned, but on which this analysis depends, is the upper wavenumber used in the analysis, $k_{\rm max}$. Linear theory is reasonably robust up to $k = 0.2h{\rm Mpc}^{-1}$ and so this is set as the upper limit here. In practice it was found that the optimal survey parameters changed little with changes in $k_{\rm max}$, so that these results with $k_{\rm max} = 0.2h{\rm Mpc}^{-1}$ are quite general.

The results show that the optimal strategy is to be as accurate as possible in the distance determinations at the expense of going deep. In fact the absolute optimal strategy is found to be

impossible with current distance indicator limits – if the distances could be better constrained then we could do even better. The reason for this is that by $k \approx 0.2h\mathrm{Mpc}^{-1}$ the power spectrum is already becoming dominated by shot noise. If there were more sources (and so less shot noise) this would have the same effect as increasing α_v in the degenerate parameter η and so the optimal value for σ_0 would also increase. The bottom line is that the survey should be as well sampled as possible, and its distance indicators should be as accurate as possible, before going deep. The expected precision of the 6dFGS distance indicators is around 20%.

3.7.4 Cosmological parameter forecasts from the velocity survey

In the linear regime the properties of the velocity field can be specified by just two parameters – the amplitude of the velocity power spectrum, A_v, and the shape parameter, Γ. Figure 17 shows the fractional error expected on A_v and Γ for an optimal 6dF velocity survey, as a function of maximum wavenumber analyzed. If the analysis is truncated at $k = 0.2\,h^{-1}\mathrm{Mpc}$, the fractional uncertainty in both A_v and Γ is expected to be $\approx 25\%$, ignoring any systematic uncertainties.

Figure 17 also shows the parameter covariance coefficient as a function of wavenumber. The value is roughly constant at:

$$\gamma_{A_v\Gamma}(k) = \frac{\langle \Delta A_v \Delta\Gamma \rangle}{\Delta A_v \Delta\Gamma} \approx -0.8. \tag{147}$$

As with the galaxy power spectrum and for the same reasons, the amplitude and shape are strongly anti–correlated. It is because of this anticorrelation that a joint constaint of A_v and Γ is difficult. If we assume a value for Γ obtained from the redshift survey, Figure 16 shows that A_v may be constained on its own to $\approx 6\%$. Figure 18 shows the source of this improvement. The strong anticorrelation between A_v and Γ makes the marginalised uncertainties on both parameters very large, but the conditional uncertainty on the amplitude alone, is small.

3.8 Combining the density field with the velocity field

To use the redshift and velocity data sets from the 6dFGS to their full advantage it will be necessary to consider their combination. In the sense that the velocity field traces underlying dark

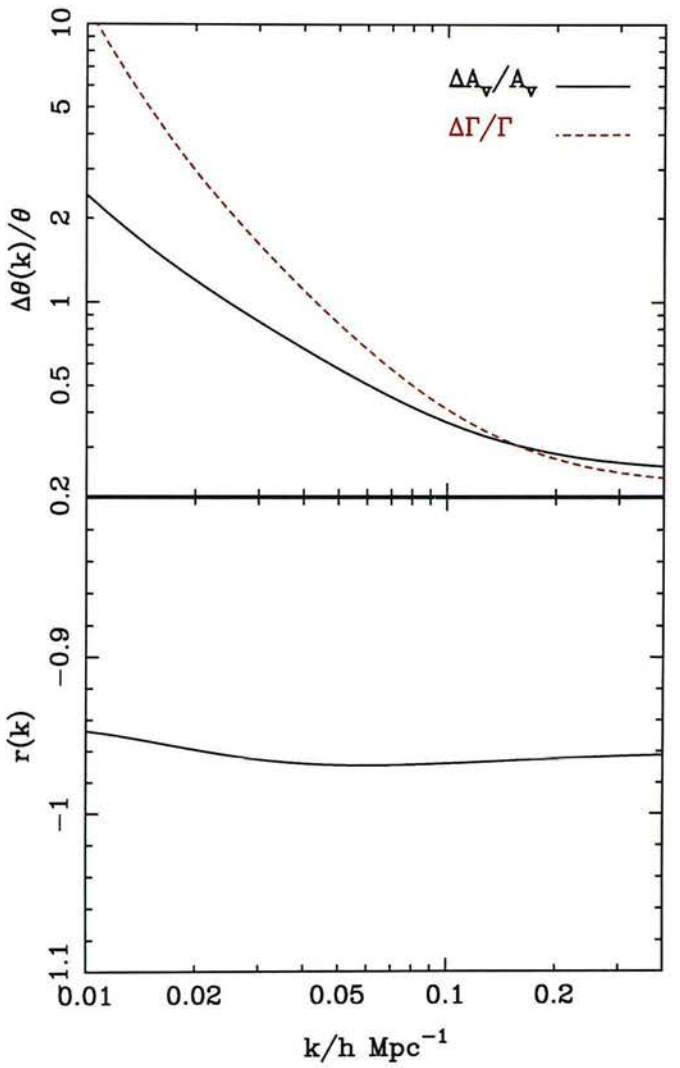

Figure 17: *The expected fractional error (top) and correlation coefficient (bottom) for the amplitude and shape parameter of the 6dF velocity power spectrum. The survey parameters are discussed in the text.*

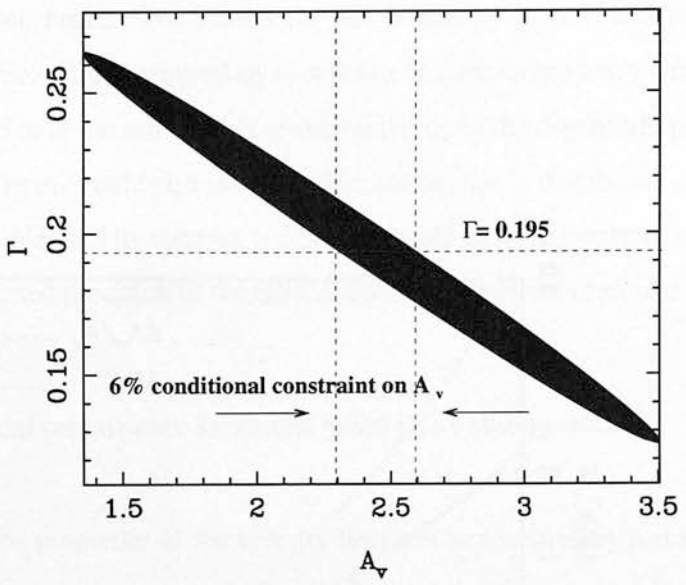

Figure 18: *The expected error ellipse on A_v and Γ from the velocity survey. If a prior value for Γ is assumed – perhaps based on the value constrained from the redshift survey – then a $\sim 6\%$ conditional constraint for A_v can be found.*

matter, and galaxies represent only the luminous fraction of the mass budget, there is clearly going to be information in a velocity–redshift data set concerning the relationship between the two. The present section concerns the construction of the bivariate Fisher matrix – i.e. the information matrix of the two data sets together. Presented in the next section are the results of using this Fisher 'supermatrix'.

3.8.1 Constructing the 'super–matrix'

The definition of the data vector in any analysis is at the discretion of the person performing it. However, the information content of a given data set should not depend on the method chosen to analyse it. Below, two methods – differing on the definition of the data – are considered, and the Fisher matrix for each is derived. Happily the final expressions for the two are identical.

The starting point for both 'mode–mode' analysis and 'radical compression' is the general expression for the Fisher matrix of a discrete data set $\mathbf{x}$ with mean $\boldsymbol{\mu}$:

$$\mathbf{F}_{ij} = \frac{1}{2}\mathrm{Tr}\left[\mathbf{C}^{-1}\frac{\partial \mathbf{C}}{\partial \theta_{\mathrm{i}}}\mathbf{C}^{-1}\frac{\partial \mathbf{C}}{\partial \theta_{\mathrm{j}}} + \mathbf{C}^{-1}\left(\frac{\partial \boldsymbol{\mu}}{\partial \theta_{\mathrm{i}}}\frac{\partial \boldsymbol{\mu}^{\mathrm{t}}}{\partial \theta_{\mathrm{j}}} + \frac{\partial \boldsymbol{\mu}}{\partial \theta_{\mathrm{j}}}\frac{\partial \boldsymbol{\mu}^{\mathrm{t}}}{\partial \theta_{\mathrm{i}}}\right)\right]. \qquad (148)$$

(TTH, and equation (220) in the Appendix) where $\mathbf{C}$ is the data covariance matrix

$$\mathbf{C} = \langle (\mathbf{x} - \boldsymbol{\mu})(\mathbf{x} - \boldsymbol{\mu})^t \rangle. \tag{149}$$

It is easiest to begin in the limit of discrete data and move to the continuum later on.

3.8.2 Mode – mode analysis

The first method is akin to that used for the single field analysis. The data vector consists of a set of discrete Fourier modes of *both* the overdensity field *and* the radial gradient of the peculiar velocity field:

$$\mathbf{x} = (\delta(k_1), ...\delta(k_N), u'(k_1), ...u'(k_N)). \tag{150}$$

This time our data covariance matrix is much larger and involves non–zero off–diagonal elements:

$$\mathbf{C} = \begin{pmatrix} \langle \delta(k_1)\delta(k_1) \rangle & \langle \delta(k_1)\delta(k_2) \rangle & ..... & \langle \delta(k_1)u'(k_1) \rangle & \langle \delta(k_1)u'(k_2) \rangle & ..... \\ \langle \delta(k_2)\delta(k_1) \rangle & \langle \delta(k_2)\delta(k_2) \rangle & ..... & \langle \delta(k_2)u'(k_1) \rangle & \langle \delta(k_2)u'(k_2) \rangle & ..... \\ ..... & ..... & ..... & ..... & ..... & ..... \\ \langle u'(k_1)\delta(k_1) \rangle & \langle u'(k_1)\delta(k_2) \rangle & ..... & \langle u'(k_1)u'(k_1) \rangle & \langle u'(k_1)u'(k_2) \rangle & ..... \\ \langle u'(k_2)\delta(k_1) \rangle & \langle u'(k_2)\delta(k_2) \rangle & ..... & \langle u'(k_2)u'(k_1) \rangle & \langle u'(k_2)u'(k_2) \rangle & ..... \\ ..... & ..... & ..... & ..... & ..... & ..... \end{pmatrix} \tag{151}$$

We assume that modes of different wavelength are uncorrelated and so have zero covariance, however terms such as $\langle \delta(k_1)u'(k_1) \rangle$, $\langle \delta(k_2)u'(k_2) \rangle$ etc. are non–zero so the covariance matrix looks something like:

$$\mathbf{C} = \begin{pmatrix} \langle \delta(k_1)\delta(k_1) \rangle & 0 & ..... & \langle \delta(k_1)u'(k_1) \rangle & 0 & ..... \\ 0 & \langle \delta(k_2)\delta(k_2) \rangle & ..... & 0 & \langle \delta(k_2)u'(k_2) \rangle & ..... \\ ..... & ..... & ..... & ..... & ..... & ..... \\ \langle u'(k_1)\delta(k_1) \rangle & 0 & ..... & \langle u'(k_1)u'(k_1) \rangle & 0 & ..... \\ 0 & \langle u'(k_2)\delta(k_2) \rangle & ..... & 0 & \langle u'(k_2)u'(k_2) \rangle & ..... \\ ..... & ..... & ..... & ..... & ..... & ..... \end{pmatrix} \tag{152}$$

This can be written in terms of four separate (diagonal) covariance matrices $\mathbf{C}_{11}$, $\mathbf{C}_{22}$ and $\mathbf{C}_{12} = \mathbf{C}_{21}$, where

$$\mathbf{C} = \begin{pmatrix} \mathbf{C}_{11} & \mathbf{C}_{12} \\ \mathbf{C}_{21} & \mathbf{C}_{22} \end{pmatrix} \tag{153}$$

and $\mathbf{C}_{11} = \mathrm{Cov}(\delta_\mathrm{i}\delta_\mathrm{j})$, $\mathbf{C}_{22} = \mathrm{Cov}(\mathrm{u}'_\mathrm{i}\mathrm{u}'_\mathrm{j})$ and $\mathbf{C}_{12} = \mathrm{Cov}(\delta_\mathrm{i}\mathrm{u}'_\mathrm{j})$. Substitution of this into equation (148) gives

$$\mathbf{F}_{ij} = \frac{1}{2}\mathrm{Tr}\left[\mathbf{C}^{-1}\frac{\partial \mathbf{C}}{\partial \theta_\mathrm{i}}\mathbf{C}^{-1}\frac{\partial \mathbf{C}}{\partial \theta_\mathrm{j}}\right] \tag{154}$$

since the data set has zero mean. Moving to the continuum limit we obtain equation (101) which when written out fully gives:

$$\begin{aligned}\mathbf{F}_{ij} = \frac{1}{2}\int d^3r \frac{d^3k}{(2\pi)^3}(&C_{22}^2\partial_i C_{11}\partial_j C_{11} + 2C_{11}C_{22}\partial_i C_{12}\partial_j C_{12} + C_{11}^2\partial_i C_{22}\partial_j C_{22} \\ &-2C_{22}C_{12}[\partial_i C_{12}\partial_j C_{11} + \partial_i C_{11}\partial_j C_{12}] - 2C_{11}C_{12}[\partial_i C_{12}\partial_j C_{22} + \partial_i C_{22}\partial_j C_{12}] \\ &+C_{12}^2[\partial_i C_{22}\partial_j C_{11} + 2\partial_i C_{12}\partial_j C_{12} + \partial_i C_{11}\partial_j C_{22}])/(2(C_{11}C_{22} - C_{12}^2)^2)\end{aligned} \tag{155}$$

where

$$\begin{aligned} C_{11} &= P_{gg}(k) + N_{gg}(k), \\ C_{22} &= P_{u'u'}(k) + N_{u'u'}(k), \\ C_{12} &= P_{u'g}(k) = r_g\sqrt{P_{gg}(k)P_{u'u'}(k)}. \end{aligned} \tag{156}$$

3.8.3 Radical Compression

An alternative to the above procedure is to choose the data set whose mean can be written in terms of the three power spectra. Thus we define a data set $\mathbf{x}$ whose components are all possible non zero values of $\delta(\mathbf{k})\delta^*(\mathbf{k}')\delta_D(\mathbf{k}-\mathbf{k}')$, $u'(\mathbf{k})u'^*(\mathbf{k}')\delta_D(\mathbf{k}-\mathbf{k}')$ and $\delta(\mathbf{k})u'^*(\mathbf{k}')\delta_D(\mathbf{k}-\mathbf{k}')$ where δ_D is the Dirac delta function. As in the previous case it is easiest to start in the discrete limit and move to the continuum. The data vector $\mathbf{x}$ is therefore defined as all possible values of the data set for a finite set of Fourier modes $\mathbf{k}$:

$$\begin{aligned}\mathbf{x} = &(\delta(k_1)\delta^*(k_1), ...\delta(k_N)\delta^*(k_N), u'(k_1)u'^*(k_1), ...u'(k_N)u'^*(k_N), \\ &\delta(k_1)u'^*(k_1), ..., \delta(k_N)u'^*(k_N)).\end{aligned} \tag{157}$$

The covariance matrix for this data vector is larger than that in the previous section, but its components may be simplified considerably. The covariance of two variates x_1 and x_2 is

simply

$$\mathrm{Cov}(\mathrm{x}_1, \mathrm{x}_2) = \langle \mathrm{x}_1 \mathrm{x}_2 \rangle - \langle \mathrm{x}_1 \rangle \langle \mathrm{x}_2 \rangle \tag{158}$$

and when the variates x_i are themselves the products of Gaussian variates y_j this can be written

$$\begin{aligned} \mathrm{Cov}(\mathrm{x}_1, \mathrm{x}_2) &\doteq \mathrm{Cov}(\mathrm{y}_1\mathrm{y}_2, \mathrm{y}_3\mathrm{y}_4) \\ &= \langle y_1 y_2 y_3 y_4 \rangle - \langle y_1 y_2 \rangle \langle y_3 y_4 \rangle \\ &= \langle y_1 y_2 \rangle \langle y_3 y_4 \rangle + \langle y_1 y_3 \rangle \langle y_2 y_4 \rangle + \langle y_1 y_4 \rangle \langle y_2 y_3 \rangle - \langle y_1 y_2 \rangle \langle y_3 y_4 \rangle \\ &= \langle y_1 y_3 \rangle \langle y_2 y_4 \rangle + \langle y_1 y_4 \rangle \langle y_2 y_3 \rangle. \end{aligned} \tag{159}$$

From this expression the key elements of the data covariance matrix can be found. For any general wavenumber:

$$\begin{aligned} \mathrm{Cov}(\delta\delta^*, \delta\delta^*) &= 2\langle \delta\delta^* \rangle^2 \\ \mathrm{Cov}(\mathrm{u}'\mathrm{u}'^*, \mathrm{u}'\mathrm{u}'^*) &= 2\langle \mathrm{u}'\mathrm{u}'^* \rangle^2 \\ \mathrm{Cov}(\delta\mathrm{u}'^*, \delta\mathrm{u}'^*) &= \langle \delta\mathrm{u}'^* \rangle^2 + \langle \delta\delta^* \rangle \langle \mathrm{u}'\mathrm{u}'^* \rangle \end{aligned} \tag{160}$$

representing the diagonal components, and

$$\begin{aligned} \mathrm{Cov}(\delta\delta^*, \mathrm{u}'\mathrm{u}'^*) &= 2\langle \delta\mathrm{u}'^* \rangle^2 \\ \mathrm{Cov}(\delta\delta^*, \delta\mathrm{u}'^*) &= 2\langle \delta\delta^* \rangle \langle \delta\mathrm{u}'^* \rangle \\ \mathrm{Cov}(\mathrm{u}'\mathrm{u}'^*, \delta\mathrm{u}'^*) &= 2\langle \mathrm{u}'\mathrm{u}'^* \rangle \langle \delta\mathrm{u}'^* \rangle \end{aligned} \tag{161}$$

representing the off–diagonal components. From equations (160) and (161) we can rewrite the data covariance matrix in terms of smaller constituent, diagonal matrices in analogy with equation (153):

$$\mathbf{C} = 2 \begin{pmatrix} \mathbf{C}'_{11} & \mathbf{C}'_{12} & \mathbf{C}'_{13} \\ \mathbf{C}'_{21} & \mathbf{C}'_{22} & \mathbf{C}'_{23} \\ \mathbf{C}'_{31} & \mathbf{C}'_{32} & \frac{1}{2}\mathbf{C}'_{33} \end{pmatrix} \tag{162}$$

where this time the constituent matrices are defined as

$$\mathbf{C}'_{11} = \langle \delta(k_i) \delta^*(k_j) \rangle^2 \tag{163}$$

$$\mathbf{C}'_{22} = \langle u'(k_i) u'^*(k_j) \rangle^2 \tag{164}$$

$$\mathbf{C}'_{33} = \langle \delta(k_i) u'^*(k_j) \rangle^2 + \langle \delta(k_i) \delta^*(k_j) \rangle \langle u'(k_i) u'^*(k_j) \rangle \tag{165}$$

for the diagonal elements, and:

$$\begin{aligned} \mathbf{C}'_{12} &= \langle \delta(k_i) u'^*(k_j) \rangle^2 \\ \mathbf{C}'_{13} &= \langle \delta(k_i) \delta^*(k_j) \rangle \langle \delta(k_i) u'^*(k_j) \rangle \\ \mathbf{C}'_{23} &= \langle u'(k_i) u'^*(k_j) \rangle \langle \delta(k_i) u'^*(k_j) \rangle \end{aligned} \tag{166}$$

for the off–diagonal elements. In a similar way the vector for the mean can be written in terms of three constituent vectors:

$$\mu_i = (P_{gg}(k_i), P_{u'u'}(k_i), P_{gu'}(k_i)). \tag{167}$$

Substituting equations (162) through to (167) into equation (148) and moving to the continuum limit we get back to the same expression for the Fisher super–matrix (equation (155)).

3.9 Parameter forecasts for the combined data sets

3.9.1 Three parameter set

To see how combining redshift and velocity data affects parameter estimation we begin with our three parameter set; A_g, Γ and β. Figure 19 shows the uncertainties and correlations of this set for the combined 6dF redshift and velocity surveys. Comparing with just redshift information, in Figure 12 we see that the constraints on A_g and Γ remain largely unchanged and there is a slight improvement in the constraint on β. So for this three parameter set there is relatively little improvement. This is because the inclusion of the velocity data does nothing to break the main degeneracy between A_g and Γ. The other parameters were relatively uncorrelated. The lower panel in Figure 19 shows that the correlations remain unchanged from the redshifts–only analysis.

Figure 20 shows the predicted $\Delta \log(L) = -1/2$ likelihood contours about the fiducial maximum likelihood point in a two–parameter space. These contours represent the likelihood marginalised in each case over the one remaining parameter. The outer contour is that from the redshift survey alone whereas the inner hatched region is the constraint when velocity data is included. The maximum wavenumber analysed here is $k = 0.2\,h^{-1}\mathrm{Mpc}$.

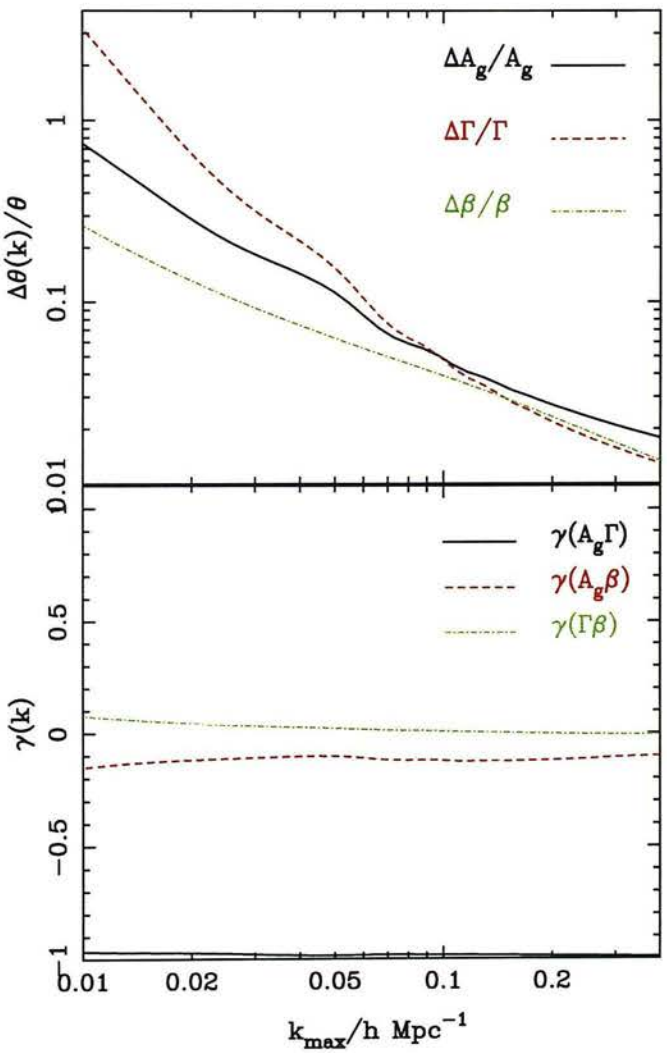

Figure 19: *Supermatrix fractional errors and correlations between A_g, Γ and β. These predictions are for a combination of the 6dF redshift and peculiar velocity surveys.*

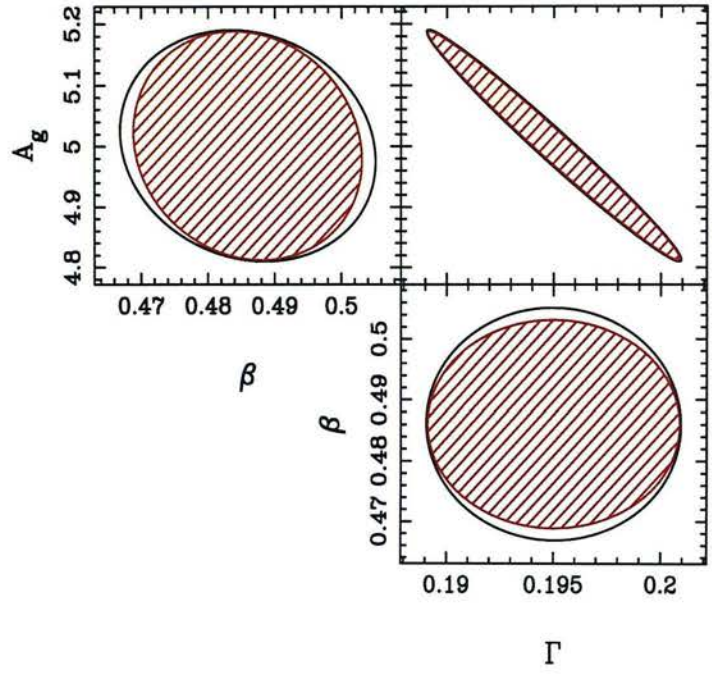

Figure 20: 1σ *contours for likelihood marginalised over one free parameter. The outer contour is for the 6dF redshift survey only. Hatched regions indicate the constraint from combining with the peculiar velocity survey.*

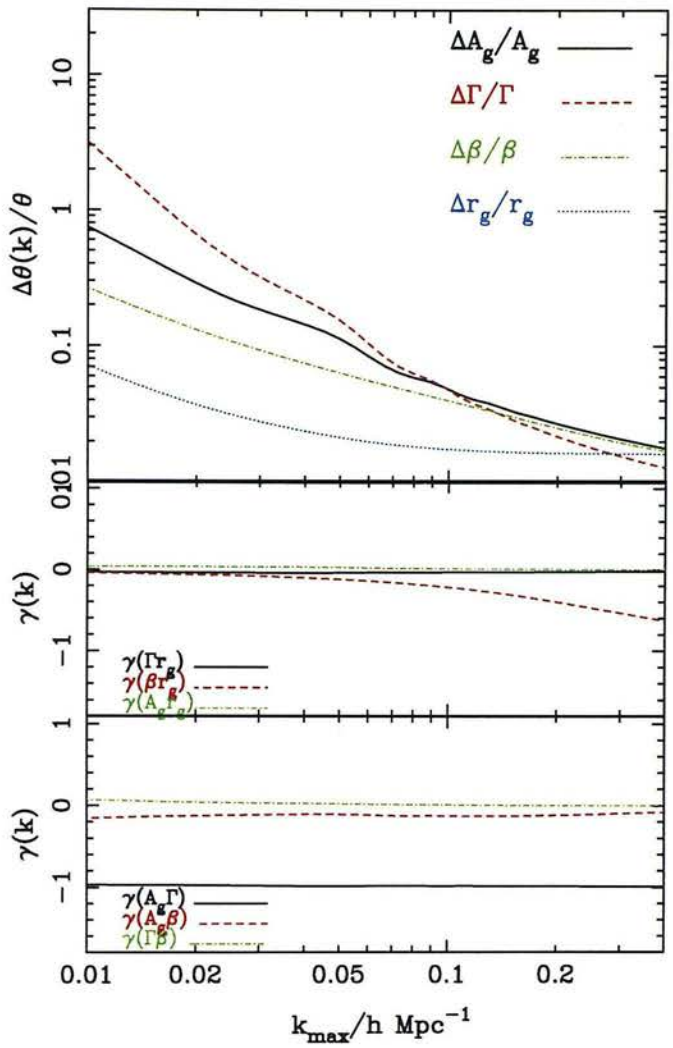

Figure 21: *Supermatrix fractional errors and correlations between A_g, Γ, β and r_g.*

3.9.2 Four parameter set

As with the galaxy survey, we can add in a new parameter and study its effect on a likelihood analysis. Adding in ω_b has little effect on A_g and β, but degrades Γ, as was the case for the redshift survey alone. The baryonic density is not detected from the combined survey anyway, nor is the baryon fraction. We shall not consider this set further.

Figure 21 shows the marginalised uncertainties of a four parameter set including the galaxy–mass correlation parameter, r_g. Comparing with Figure 19 we see that A_g and Γ are unchanged by the inclusion of the new parameter, since these are practically independent of r_g. The red-

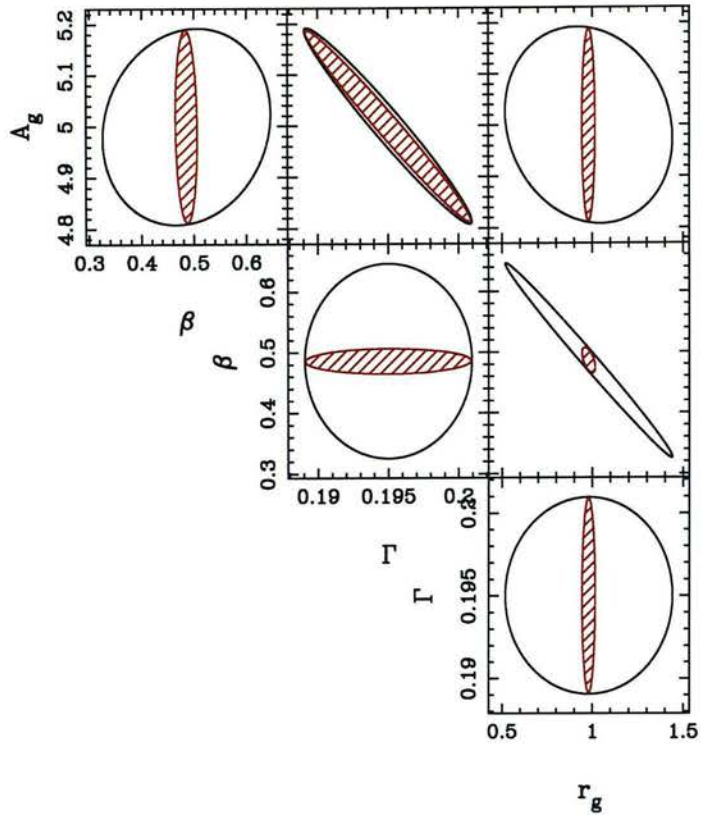

Figure 22: 1σ *contours for likelihood marginalised over two free parameters.*

shift distortion parameter β, has degraded slightly, due to its relatively weak correlation with r_g, but is still measurable at the 3% level. The big change from the redshifts–only analysis is in the galaxy–mass correlation parameter r_g, which is now measurable from the combined surveys, with a formal uncertainty of better than 2%. Comparing Figure 21 with its redshift–survey–only equivalent, Figure 14, it is clear that there has been a great improvement in the joint constraint of β and r_g. The reason for this can be seen by comparing the correlations between the two parameters in the lower panels of both figures. In the redshifts–only analysis β and r_g are strongly anticorrelated whereas when the redshifts are combined with peculiar velocities this degeneracy is broken and the parameters become almost independent.

Again we have plotted the expected likelihood contours for this analysis in order to help visualise the improvement in the constraint. In Figure 22 the contours represent the likelihood after marginalization over the two remaining parameters in a four parameter fit. Again the hatched region shows the improvement from combining the data sets. Particularly interesting

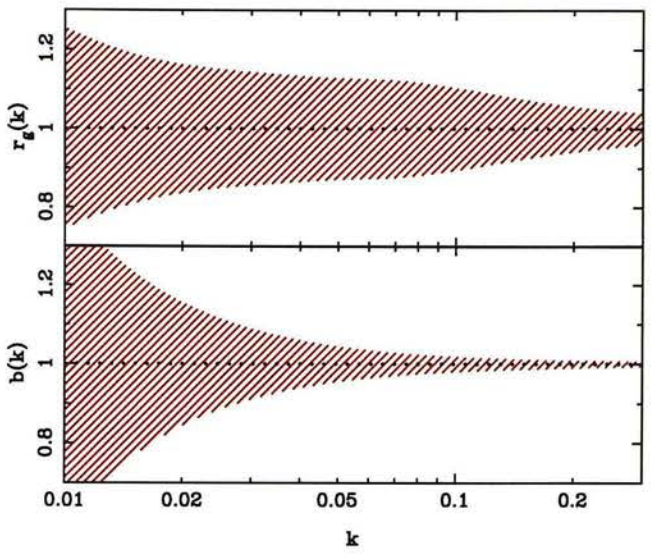

Figure 23: *Constraining scale dependence:* $\Delta b(k)$ *and* $\Delta r_g(k)$.

is the panel showing contours for β and r_g which illustrates the breaking of the degeneracy between the two, and the subsequent improvement in their joint constraint.

3.9.3 Scale dependence of b and r_g

Given the tight constraint on r_g and β it is interesting to consider a possible scale dependence in either r_g or the biassing parameter b. Although we have approximated b to be constant over the scale of interest, we should expect it to have some scale dependence. Another way to consider scale dependence in the biassing is to fix b as constant and allow the parameter r_g to vary with k. The upper panel of Figure 23 shows the expected error bars on a measure of r_g band averaged over a logarithmic passband of width $\Delta \ln k$. The width $\Delta \ln k$ is marked on the plot as a series of dots. The simple biassing parameter b is fixed at $b = 1$ for this analysis. The results are extremely encouraging; $r_g(k)$ can be measured to within about 10% over a wide range of scales improving to 4.5% on the smallest scales. The lower panel shows a similar error bar prediction for estimates of the band–averaged $b(k)$. This plot assumes r_g to be fixed at $r_g = 1$ over all scales. The bias parameter $b(k)$ is even more tightly constrained at a few percent over a wide range.

3.10 Conclusions

This chapter has presented the formalism for the individual and combined Fisher information analysis for galaxy redshift and velocity field surveys. This analysis allows us to optimise both surveys to maximise the information content for cosmological parameters, providing an estimate of the uncertainty on the measurement of the matter and velocity power spectra and the set of cosmological parameters, $(A_g, A_v, \Gamma, \beta, \omega_b, r_g)$. For both 6dF redshift and velocity surveys we find the optimal design to be as wide as possible – a result which was previously well known. In the case of the velocity survey we find the best design to be as well sampled and as accurate as possible and in the redshift survey we find an optimal sampling of around 16% but that parameter estimation is not particularly sensitive to this design parameter. We expect to be able to constrain A_g, Γ and β to around $2-3\%$ from the redshift survey. The amplitude of mass clustering Q is measurable to $\sim 5\%$. From the velocity survey A_v can be constrained to 6% but a joint constraint of A_v and Γ will have marginalised uncertainties of 25%.

We find that the major benefits of the 6dFGS are found when the velocity and redshift surveys are combined and when we wish to jointly constrain the parameters r_g and β. The parameters' degeneracy is broken when the power spectra are combined and the parameter r_g can be measured much more accurately than in any of the above surveys with just redshifts. All four parameters – A_g, Γ, β and r_g – can be measured to a few percent. Finally, the scale dependence of r_g and b can be measured with the combined data set – which at least will give credance to some of the assumptions commonly made about biassing. Clearly the great benefit of peculiar velocity information is that it tells us about the underlying mass and by combining this information with galaxy redshifts we can learn much about the relationship between luminous and dark matter.

4 Spherical Harmonics Analysis of Galaxy Surveys

4.1 Introduction

Until now peculiar velocity surveys have been small and many analyses have involved the compilation of different data sets with all the inherent uncertainties about different systematic assumptions that this entails (see e.g. Strauss and Willick 1995). The 6dFGS is the first large combined redshift and peculiar velocity survey of the local Universe with $\sim$ 170000 redshifts and $\sim$ 15000 peculiar velocities provided by one instrument and using one selection function. This provides the opportunity to combine the velocity and redshift data sets, and test directly the relationship between dark and luminous matter. Chapter 3 shows that the 6dFGS can tightly constrain the galaxy amplitude and the shape parameter, but that its great improvement on previous parameter estimation is in the joint constraint of the redshift distortion parameter β and the matter–galaxy correlation coefficient r_g.

This chapter concerns the development of a maximum likelihood technique. Simulations of the survey are used to demonstrate its ability to constrain these two parameters. Since an important result of Chapter 3 was that an all sky survey is always preferable to half–sky, the technique is also tested on a suite of all sky simulations using the same 6dF radial selection criteria. The method involves the decomposition of the galaxy overdensity field and radial peculiar velocity field into a set of basis functions chosen to reflect the spherical symmetry of the problem. Spherical Bessel functions are used in the radial direction and spherical harmonics in the angular directions. The method follows that set out by Heavens and Taylor (HT, 1995) and used subsequently in Ballinger et al. (2000), Tadros et al. (1999), Taylor et al (2001) and Percival et al. (P04, 2004). These papers dealt only with the density field observed in a galaxy redshift survey. Here the analysis is extended to peculiar velocity surveys and the combination of velocities and redshifts.

There are many benefits to using spherical basis functions to describe the observed fields. For one thing they allow us to separate out the angular survey window function from the radial

selection function. Also redshift space distortions only affect the radial positions of observed sources and so have no effect on the angular functions. All of this means that the angular and radial parts can be calculated separately in programs running in parallel, saving time. There are also many advantages to working in Fourier space. The underlying Fourier modes are statistically independent by virtue of the Cosmological principle. The observed modes are convolved with the survey mask and so lose this independence. However, as we increase the survey size the effect of this aliasing between modes decreases. In Fourier space operators such as the gradient become simple multiplicative factors. As well as this, different physical processes operating on different scales can be separated out in Fourier space. There is therefore a clear division between modes which are within the linear regime and those which are not.

One important point to be made about the following analysis, is that its true value is found when applied to an all–sky survey. When applied to the partial sky coverage of the present 6dF survey, the analysis is complicated by convolution with the survey mask which introduces practical problems in the analysis. As the results towards the end of this Chapter show, an all–sky survey would not only simplify the analysis considerably, but also significantly improve the constraints on parameters.

The chapter is set out as follows. In Section 4.3 the model is outlined. Section 4.4 shows how to transform the radial velocities to provide the data set that has been modelled. In Section 4.5 a model for the noise expected from both the density and velocity modes is found. Section 4.6 describes the likelihood analysis and prior assumptions. In Section 4.7 simulations of the 6dFGS are constructed. Section 4.8 discusses ways to optimally compress the data to speed up the analysis, and in Section 4.9 some practical issues are discussed. Finally Section 4.10 shows the results of the analysis tested on the simulations. There now follows a brief review of the history of this type of analysis.

4.2 A brief review of spherical harmonics analyses

The use of a spherical harmonics formalism for extracting information from density fields was proposed by Fisher et al. (1994b). They used this formalism to estimate β from redshift space distortions in the 1.2-Jy IRAS redshift survey. The same formalism was adopted in Fisher et al.

(1995b) as a means of recovering the real space density, velocity and potential fields from the same survey. The real space density harmonics were recovered from the redshift space modes by inverting the coupling matrix used in Fisher et al. (1994b) to calculate the likelihood, and using a Wiener filter to attenuate the effect of shot noise in the inversion. The velocity and potential modes were then shown to follow from simple linear transformations of the density modes. This paper also dealt with the question of suitable boundary conditions, which limit the analysis to a discrete set of possible wavenumbers. The formalism to this point had been developed under the simplifying assumptions of an all sky survey with a perfectly uniform mask. Heavens and Taylor (1995) expanded the formalism to describe the convolution of the modes with general survey masks (and the subsequent aliasing of power). They also introduced a further radial scattering matrix to model the distorting effects of small-scale velocity structure. Both methods were subsequently applied to the PSCz catalogue. Tadros et al. (1999) used the likelihood analysis of Heavens and Taylor (1995) to constrain β, the shape parameter and the amplitude of mass fluctuations. Taylor et al (2001) and Ballinger et al. (2000) discussed data compression with respect to this analysis. Schmoldt et al. (1999) used the Wiener reconstruction method of Fisher et al. (1995b) to obtain the density and velocity fields and β from the PSCz survey. More recently, Percival et al. (2004) applied a spherical harmonics likelihood method to the 2dF Galaxy Survey, constraining β marginalised over the power spectrum shape, and using an updated treatment of luminosity dependent bias. The latest step, taken in this Chapter, is to extend the formalism to a joint likelihood analysis of velocity and density modes, where the velocity modes are directly measured from a peculiar velocity survey. As discussed in the previous Chapter, a joint analysis such as this, makes full use of the parameter-constraining potential of the 6dFGS.

4.3 Modelling the fields

4.3.1 The density field

The density field is modelled as outlined in HT. The redshift space density field may be expanded in a series of spherical basis functions such that:

$$\rho(s) = \sum_{lmn} c_{ln} \rho^s_{lmn} j_l(k_{ln} s) Y_{lm}(\theta, \phi) \tag{168}$$

where the superscript 's' denotes a redshift space quantity. The functions j_l and Y_{lm} are spherical Bessel functions and spherical harmonics respectively. They are defined in Appendix B.1 along with their respective orthogonality relations and the normalisation constant c_{ln}. The orthogonality relations can be used to obtain the coefficients of the expansion ρ^s_{lmn} as:

$$\rho^s_{lmn} = c_{ln} \int d^3s\, w(s)\rho(s) j_l(k_{ln}s) Y^*_{lm}(\theta, \phi). \tag{169}$$

The data vector is defined as $D_\mu \equiv (\rho^s_{lmn} - (\rho_0)_{lmn})$ where ρ_0 is the unperturbed density field, synonomous with the selection function of the survey. For convenience μ is used as a shorthand notation for the modes (l, m, n) and likewise ν for (l', m', n'). After perturbing the selection function to linear order to account for the fact that measurements are made in redshift space, the observable density perturbations may be written in matrix form:

$$\boldsymbol{D} = b(\boldsymbol{\Phi} + \beta \mathbf{V})\boldsymbol{\delta}. \tag{170}$$

The matrices Φ^ν_μ and V^ν_μ are defined in the Appendix B.2 where equation (170) is explicitly proven. The choice of weighting function is also addressed here. The redshift distortion parameter β is the ratio of a function of the density parameter $f(\Omega_m)$ and the biassing parameter b discussed in the next section. An additional distortion to be modelled is that due to small scale scattering from virialised peculiar velocities. This is addressed in Appendix B.3.

4.3.2 Stochastic Biassing

In previous spherical harmonics analyses, it has been necessary to assume a relation for the biassing between the matter density field and the galaxy distribution. Although well founded simple relations exist, they disguise our ignorance about the complexities of galaxy formation. It is well known that a single biassing factor cannot be used for all galaxies since galaxies of different types cluster differently. P04 use the luminosity dependent biassing scheme advocated by Norberg et al. (2001). In fact local factors such as the temperature of the Inter–Galactic Medium or the presence of magnetic fields must also play a part in the strength of clustering. It is therefore convenient to model the physics of galaxy biassing by describing it as having a random stochastic error on top of the simple linear biassing relation (Dekel and Lahav 1998). The relationship between the fields may be written:

$$\delta_{gal}(k_\mu) = b_L \delta_{mass}(k_\mu) + \epsilon \tag{171}$$

where ϵ is a random factor which accounts for all possible effects in the relationship between mass and galaxies which are not dealt with by a simple linear dependence. We have $\langle\epsilon\rangle = 0$ and we define $\langle\epsilon^2\rangle = \sigma_\epsilon^2$. It is now possible to define a biassing factor

$$b^2 \equiv \frac{\langle\delta_g\delta_g\rangle}{\langle\delta_m\delta_m\rangle} = b_L^2 + \frac{\sigma_\epsilon^2}{\langle\delta_m\delta_m\rangle} \tag{172}$$

where b_L is a constant. So there is clearly some scale dependence in b, however it is convenient to approximate this as constant over the scales of interest (which is true if the unknown factor σ_ϵ^2 is small). A correlation parameter between the density and mass can be defined as

$$r_g = \frac{\langle\delta_g\delta_m\rangle}{\sqrt{\langle\delta_g\delta_g\rangle\langle\delta_m\delta_m\rangle}} = \frac{b_L}{b}. \tag{173}$$

In Chapter 3 it was shown that the 6dF redshift survey on its own does not allow the joint constraint of r_g and β because of the strong correlation between the two parameters. However, the addition of peculiar velocities breaks this degeneracy. Whereas in HT and subsequent papers, r_g is simply set to one, it can now be left as a free parameter and constrained in the likelihood analysis. The galaxy–galaxy covariance terms become:

$$\langle \boldsymbol{D}\boldsymbol{D}^t\rangle = b^2\left[\boldsymbol{\Phi}\mathbf{P}\boldsymbol{\Phi}^t + \beta r_g(\boldsymbol{\Phi}\mathbf{P}\mathbf{V}^t + \mathbf{V}\mathbf{P}\boldsymbol{\Phi}^t) + \beta^2\mathbf{V}\mathbf{P}\mathbf{V}^t\right]. \tag{174}$$

where the diagonal matrix $\mathbf{P}$ is the underlying matter power spectrum. The power spectrum is modelled using the transfer function of Eisenstein and Hu (1998) discussed in Section 1.11.5.

4.3.3 The velocity field

It can be shown that to first order the transform of the radial velocity or the radial velocity gradient is the same in real space as in redshift space. That is to say that linear redshift distortions may be neglected in the velocity field analysis. We wish to work with the radial gradient of the radial velocity field as discussed in Chapter 3. The transformation into spherical harmonics and spherical Bessel functions provides the modes:

$$u'_\mu = c_\mu \int d^3r\, w(r)\rho(r)\frac{du(\mathbf{r})}{dr} j_\mu(k_\mu r) M(\Omega) Y_\mu^*(\Omega) \tag{175}$$

which is the velocity field counterpart to equation (169). The dash indicates that the modes are those of the velocity gradient. For convenience of notation $\Omega = (\theta, \phi)$ denotes angular

position. The modes defined in equation (175) constitute the data set to be used in the velocity part of the likelihood analysis. As is the case when modelling the redshift space distortions for the density field, it is necessary here to use the dynamical relation between velocity and density fields:

$$u(\mathbf{r}) = \Omega_m^{0.6} \sum_\nu c_\nu k_\nu^{-2} \delta_\nu \frac{dj_\nu(k_\nu r)}{dr} Y_\nu(\Omega) \tag{176}$$

(Fisher et al. (1994c), HT and Appendix B.2). It is then a simple matter to differentiate the Bessel function a second time to obtain the radial velocity gradient. Substituting this into equation (175) gives an expression for the velocity modes in terms of the underlying density:

$$u'_\mu = \Omega_m^{0.6} \Psi_\mu^\nu \delta_\nu, \tag{177}$$

(summation implied) where

$$\Psi_\mu^\nu \equiv \Psi_\nu^\mu W_\mu^\nu, \tag{178}$$

and

$$\Psi_\nu^\mu = \frac{c_\mu c_\nu}{k_\nu^2} \int r^2 dr\, w(r) \rho(r) j''_\nu(k_\nu r) j_\mu(k_\mu r), \tag{179}$$

$$W_\mu^\nu = \int d\Omega\, Y_\nu(\Omega)\, M(\Omega)\, Y_\mu^*(\Omega). \tag{180}$$

Primes denote derivatives with respect to r.

4.3.4 The Full Covariance Matrix

For an all sky survey the orthogonality of the spherical harmonics means that the angular integral in equation (180) becomes a product of delta functions $\Delta_{ll'} \times \Delta_{mm'}$. For general survey masks there is a mixing of the modes due to the window matrix which may also now be complex. In order to deal with this it is convenient to separate the observed data into real and imaginary parts and treat them as independent. This corresponds to doubling the size of matrices $\mathbf{\Phi}$, $\mathbf{\Psi}$ and $\mathbf{V}$. The full covariance matrix then takes the form:

$$\mathbf{C} = \begin{pmatrix} \langle \boldsymbol{D}\boldsymbol{D}^t \rangle & \langle \boldsymbol{u}'\boldsymbol{D}^t \rangle \\ \langle \boldsymbol{D}\boldsymbol{u}'^t \rangle & \langle \boldsymbol{u}'\boldsymbol{u}'^t \rangle \end{pmatrix} \tag{181}$$

plus noise, where the individual elements can be written:

$$\langle \boldsymbol{D}\boldsymbol{D}^t \rangle = \frac{b^2}{2} \left[\mathbf{\Phi}\mathbf{P}\mathbf{\Phi}^t + \beta r_g (\mathbf{\Phi}\mathbf{P}\mathbf{V}^t + \mathbf{V}\mathbf{P}\mathbf{\Phi}^t) + \beta^2 \mathbf{V}\mathbf{P}\mathbf{V}^t \right],$$

$$\langle \boldsymbol{u}'\boldsymbol{u}'^t \rangle = \frac{1}{2}\beta^2 b^2 \boldsymbol{\Psi}\mathbf{P}\boldsymbol{\Psi}^t \tag{182}$$

and:

$$\langle \boldsymbol{D}\boldsymbol{u}'^t \rangle = \langle \boldsymbol{u}'\boldsymbol{D}^t \rangle^t = \frac{1}{2}\beta b^2 (r_g \boldsymbol{\Phi}\mathbf{P}\boldsymbol{\Psi}^t + \beta \mathbf{V}\mathbf{P}\boldsymbol{\Psi}^t). \tag{183}$$

Note the factor of a half due to the separation of real and imaginary parts. The full covariance matrix is also explicitly set out in equation (259) in Appendix C.2.

4.4 Transforming the data

4.4.1 Transforming the radial velocity

The velocity data set as modelled in Section (4.3) is the *radial gradient* of the radial velocity. Of course the survey data only provide the radial velocities themselves. In order to transform the data, it is necessary to integrate the definition in equation (175) analytically to get the transform in terms of the radial velocity rather than its gradient. Taking the radial part of equation (175) and integrating by parts gives:

$$\begin{aligned} u'_\mu &= -c_\mu \int dr u(r) r^2 \left(\frac{2}{r} + \frac{d}{dr}\right) w(r)\rho_0(r) j_\mu(k_\mu r) \\ &\quad \times \int M(\Omega) Y^*_\mu(\Omega) d\Omega, \end{aligned} \tag{184}$$

using a boundary condition that the radial velocity falls to zero at the survey's edge. In the last step the approximation $u(r)\rho(r) \approx u(r)\rho_0(r)$ has been used, as is valid in the linear regime.

4.4.2 Correcting for the motion of the local group

So far it has been assumed that the mean radial velocity is zero. This is of course untrue if peculiar velocities are measured in the local group rest frame rather than the CMB frame. To clarify this point $u(r)$ is defined as the radial velocity as measured in the CMB frame and when transforming velocities measured in the local group one can change frames by adding a mean–field term:

$$u'_{\mu(LG)} = c_\mu \int d^3r \boldsymbol{v}_{LG}.\hat{\boldsymbol{r}} \left(\frac{2}{r} + \frac{d}{dr}\right) [w(r)\rho_0(r) j_\mu(k_\mu r)] M(\Omega) Y^*_\mu(\Omega) \tag{185}$$

where $\mathbf{v_{LG}}$ is the local group motion. Similarly for the overdensity modes, Tadros et al. (1999) show the Local Group correction to the mean field term to be:

$$(\rho_0)_{\mu(LG)} = c_\mu \int d^3r(\mathbf{v}_{LG}.\hat{\mathbf{r}})\rho_0(r)\frac{d}{dr}[w(r)j_\mu(k_\mu r)]Y^*_\mu. \tag{186}$$

As pointed out by Kaiser (1987b), an incorrect correction for the Local Group motion can lead to a spurious contribution to the derived dipole. A powerful advantage to the spherical harmonics formalism is that this so called rocket effect is contained only within the $l = 1$ mode (Fisher et al. (1995b)).

4.5 Modelling the noise

In any analysis of discrete data it is necessary to account for shot noise. In this section, the shot noise in both the density and velocity power is modelled.

4.5.1 Density field noise

The noise in the density power spectrum is stated in HT as:

$$\begin{aligned} N_{\mu\nu} = &\int \rho_0(r)w^2(r)j_\mu(k_\mu r)j_\nu(k_\nu r)r^2dr \\ &\times \int \mathcal{P}_\mu(Y^*_\mu(\Omega))M(\Omega)\mathcal{P}_\nu(Y^*_\nu(\Omega))d\Omega \end{aligned} \tag{187}$$

using the methods of Peebles (1973). Here $\mathcal{P}_{\mu,\nu}$ represent either the real or imaginary parts of the spherical harmonics depending on whether one is looking at the real or imaginary parts of the data.

4.5.2 Velocity Field noise

The derivation of the velocity field shot noise is somewhat cluttered with algebra and so is deferred to Appendix C.1. The result may be stated as:

$$N_{\mu\nu} = \int d^3r\sigma_u^2\Big[\Big(\frac{dw_\mu j_\mu}{dr} + \frac{2w_\mu j_\mu}{r}\Big)\Big(\frac{dw_\nu j_\nu}{dr} + \frac{2w_\nu j_\nu}{r}\Big)\rho_0 + \Big(\frac{dw_\mu j_\mu}{dr} + \frac{2w_\mu j_\mu}{r}\Big)w_\nu j_\nu \frac{d\rho_0}{dr} + \Big(\frac{dw_\nu j_\nu}{dr} + \frac{2w_\nu j_\nu}{r}\Big)w_\mu j_\mu \frac{d\rho_0}{dr} + w_\mu w_\nu j_\mu j_\nu \frac{d^2\rho_0}{dr^2}\Big] \times \int \mathcal{P}_\mu(Y_\mu^*(\Omega))M(\Omega)\mathcal{P}_\nu(Y_\nu^*(\Omega))d\Omega \tag{188}$$

where σ_u is the variance of the radial velocity. The notation $w_\mu \equiv w_\mu(r)$ and $j_\mu \equiv j_\mu(k_\mu r)$ is used for convenience. As discussed in Appendix B.2 the radial velocity variance in equation (188) deals with both intrinsic scatter about the Hubble relation and measurement errors.

4.6 Likelihood Analysis

It is a common assumption that density perturbations, seeded by the random phases of quantum fluctuations in the early Universe, now have a distribution of random phases. This implies that density modes are distributed as a Gaussian and that all of the statistical information of the field is contained in the variance of each mode – all higher moments being zero, and each different wavelength being independent. In reality convolution with the window function and the radial selection function, causes a mixing of modes and so a complete description of the observed field requires calculation of the full covariance matrix. Nonetheless on linear scales, maximizing the mutlivariate Gaussian likelihood will give the best possible estimate for the parameters sought. When also taking into account the exact treatment of redshift distortions to linear order, this is the most accurate possible method for analysing clustering power in the linear regime.

To perform an analysis of this type it is necessary to equate the likelihood – namely the probability of obtaining a data set given a set of model parameters – to the probability of obtaining a set of parameters given the observations. This involves Baye's theorem: $L(\mathbf{x}|\boldsymbol{\theta}) \times$

$\mathbf{p}(\boldsymbol{\theta}) = \mathbf{P}(\boldsymbol{\theta}|\mathbf{x}) \times \mathbf{p}(\mathbf{x})$ where $\mathbf{x}$ is a vector of data points and $\boldsymbol{\theta}$ a vector of model parameters. The function $p(\boldsymbol{\theta})$ is known as the prior hypothesis and $p(\mathbf{x})$ is called the evidence. In the following analysis, the likelihood is always defined as a fraction of its maximum value and so the evidence is divided out. In a likelihood analysis it is important to state the assumed prior. No prior assumptions are made concerning the parameters being constrained: all values are assumed equally likely. In fact it is not necessary to be quite so unprejudiced since r_g can only have values between -1 and 1. The prior hypothesis needed for the present analysis is a set of values for the other parameters which enter into the model power spectrum. The values $(A_g, A_v, \Gamma, \omega_b) = (5 \times 10^{-5}, \beta \times 5 \times 10^{-5}, 0.195, 0.025)$ are used.

The continuity equation implies that the velocity modes are linearly related to the density modes (equation (176)). It can also be assumed therefore, that the velocity modes are also distributed as a Gaussian and a similar analysis can be performed. It is the combination of the two however, that will provide a joint constraint of r_g and β. The likelihood functional:

$$L(\mathbf{x}|\boldsymbol{\theta}) = \frac{\mathbf{1}}{(\mathbf{2}\pi)^{\mathbf{N}/\mathbf{2}}} \frac{\mathbf{1}}{\sqrt{\det \mathbf{C}}} \exp -\frac{\mathbf{1}}{\mathbf{2}} \mathbf{x} \mathbf{C}^{-\mathbf{1}} \mathbf{x}^{\mathbf{t}} \tag{189}$$

is calculated at various points in parameter space, where the data vector $\mathbf{x} = (\delta_\mu, \mathbf{u}'_\mu)$ is comprised of both density and velocity gradient modes totalling to N data points. The model covariance matrix $\mathbf{C}$ of this data set varies from point to point in parameter space according to equations (182) and (183). The best unbiased estimate of the parameters is the vector $\boldsymbol{\theta}_{ML}$ – the point in parameter space where the likelihood peaks.

4.7 Constructing Simulations

To test the likelihood method it is first necessary to construct simulations of the eventual 6dFGS data set. The simulated catalogues are constructed from a Poisson sampling of the Hubble volume simulations. The Hubble volume simulations were carried out by the Virgo Supercomputing Consortium using computers based at the Computing Centre of the Max-Planck Society in Garching and the Edinburgh Parallel Computing Centre. The data are publicly available at http://www.mpa-garching.mpg.de/NumCos. The simulations use one billion mass particles. They provide the coordinates and velocities of test particles following an evolving gravitational potential. Twenty seven separate regions of the Hubble volume are chosen to simulate

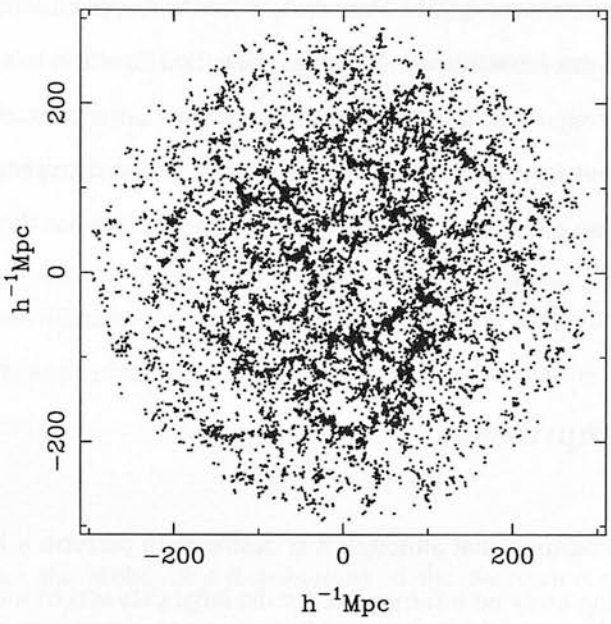

Figure 24: *Simulation of the redshift space galaxy distribution for the 6dFGS. This is a slice ranging* 30° *in declination for* 360° *of RA.*

the 6dFGS data set. Once a coordinate centre is defined it is a simple matter to calculate the proper distance of a source for the fiducial cosmology. This gives the real–space redshift albeit subject to the random scatter that produces the Fingers of God effect. The linear redshift space positions are then calculated by adding the line of sight peculiar velocity to the real space position. When assigning positions to the peculiar velocities the a priori assumption is made that their true position is given by their redshift space coordinate. This is referred to in the literature as a Type II analysis – terminology first used by Faber and Burstein (1988). As mentioned previously, to linear order it doesn't matter whether the real space or the redshift space coordinate is used. The distance indicator however, is not the *true* real space position but rather an inaccurate estimate. Inaccuracies in the measurement of proper distance can introduce vorticity into the velocity field. The redshifts are a relatively accurate measure of the *true* redshift space position and so even though they deviate systematically from the real space postions, to first order this does not matter and they are superior to the distance indicators.

Figure 24 shows a simulation of the 6dFGS in redshift space. This plot shows a 30° range in declination for 360° of RA. The survey selection function is clearly visible peaking at around $130h^{-1}$Mpc. The structure is flattened into walls by linear distortions on large scales and on smaller scales the Fingers of God effect is visible. The 6dFGS covers roughly one hemisphere

with a median redshift around 0.055. The window function applied in these rather crude simulations is just a uniform hemisphere. The radial selection function is a redshift distribution of the form $n(z) \propto z^2 \exp -(z/z_*)^{1.5}$. The galaxy distribution is normalized to around 170000 which is the expected final total. The peculiar velocities are distributed amongst a subset of around 15000 of these galaxies. The velocity selection function has the same form but cuts off at the galaxy survey's median depth of around $150h^{-1}$Mpc.

4.8 Data Compression

It has long been recognised that although it is desirable to perform a likelihood analysis on a full data set, throwing away no information, for the large data sets of modern redshift surveys it is computationally unfeasible in practice. To invert an n by n covariance matrix at N^m points in an m dimensional parameter space requires a time $t \propto n^3 \times N^m$. For this reason there has been much discussion in the literature concerning different ways to optimally compress data sets in order to preserve the most information. Here different methods are compared for the two parameter problem being discussed.

4.8.1 Signal to noise compression

The simplest method of compressing data sets is to throw away low signal to noise modes. The method consists of performing a Cholesky decomposition of the noise matrix and rotating the covariance matrix by its inverse (Press et al. 1986). Thus for a covariance matrix

$$\mathbf{C} = \mathbf{S} + \mathbf{N} \tag{190}$$

the noise is decomposed into the product $\mathbf{N} = \mathbf{L}\mathbf{L}^{\mathbf{t}}$ and the signal $\mathbf{S}$ rotated so that

$$\mathbf{C}' = \mathbf{L}^{-1}\mathbf{S}\mathbf{L}^{-\mathbf{t}} + \mathbf{I} \tag{191}$$

and the covariance matrix of the rotated data set is in the form of a signal to noise matrix plus the identity matrix. The compression then just involves diagonalizing $\mathbf{L}^{-1}\mathbf{S}\mathbf{L}^{-\mathbf{t}}$, sorting the eigenvectors in order of decreasing eigenvalue and discarding modes below a certain value. There are two drawbacks to this approach as shown in panel (a) of Figure 25. Firstly the

method is only optimal if the signal is proportional to the parameter of interest. Whereas this is true of the amplitude of the galaxy power spectrum for instance, it is not true of the distortion parameter β or the correlation coefficient r_g. Secondly this is only optimal for a conditional estimate of a single parameter rather than a marginalised estimate from a multi–parameter likelihood surface.

Figure 25 shows likelihood contours expected from the 6dF survey for r_g and β when combining redshifts and velocities. The ellipses are plotted using the equation

$$\boldsymbol{\theta}\mathbf{F}\boldsymbol{\theta}^t = \text{constant}, \tag{192}$$

which approximates the probability distributions of the parameters as being Gaussian. The constant is chosen as a certain level of confidence to which one wishes to quote results. The Fisher matrix $\mathbf{F}$ is defined as the inverse of the covariance matrix of the model parameters. For Figure 25 the Fisher matrix calculated in Chapter 3 is used. The point of plotting out this expected error ellipse is to demonstrate the correlation between the parameters, and the effects of various compression methods. Panel (a) shows a possible effect of the signal to noise compression. As the data is compressed the error on the amplitude of the signal is minimally affected as shown by the inward pointing arrows, but the error ellipse can still stretch along its major axis. The area of the ellipse will not change very much but its shape might. The result of the example in panel (a) is that conditional measurements of either parameter (where we assume a fixed value for the other) are unaffected by the compression, whereas the marginalised errors on both may be increased.

4.8.2 Karhunen–Loève eigenvalues

A more sophisticated compression method which makes use of the Fisher information matrix is the Karhunen–Loève eigenvalue compression. The Karhunen–Loève method is closely related to principal component analysis. The advantage over the signal to noise compression is that the method can be optimal for any parameter – not just amplitude terms. The Karhunen–Loève method is discussed in Tegmark et al. (TTH, 1997) and the relevant results of this paper are outlined in Appendix A.3. The optimal compression matrix for a parameter θ_i is found to be

$\mathbf{B}$ a matrix whose rows $\mathbf{b}$ are found from the eigenvectors of the equation:

$$(\mathbf{L}^{-1}\mathbf{C}_{,i}\mathbf{L}^{-t})(\mathbf{L}^{t}\mathbf{b}) = \lambda(\mathbf{L}^{t}\mathbf{b}). \quad (193)$$

The comma denotes differentiation with respect to θ_i. This compression produces a set of orthogonal eigenvectors so that no information is duplicated. The eigenvectors are then ranked according to their information content, quantified by the magnitude of the corresponding eigenvalues. A certain number are then discarded amounting to throwing away modes with negligible information on the parameter in question. The simplest way of generalising this method to a multi–parameter problem is to repeat the process for each parameter, adding the compressed data sets together.

As shown in panel (b) of Figure 25 there is still the problem that marginalised error bars can increase. The dashed lines here mark out the conditional constraints and the solid lines show the marginalised errors. The method minimises changes in the conditional uncertainty of both parameters but the ellipse can still stretch out along its major axis affecting the joint measurement of the two.

4.8.3 Generalised Optimal Mode Analysis

The solution as used in Taylor et al. (1997) and Taylor et al. (2000) is Generalised Optimal Mode Analysis (GOMA). In general a likelihood surface of any number of dimensions will have a single longest axis. This depends on both the ability of the analysis to constrain each parameter and also on the correlations between the parameters. In the two parameter case here, there is a small anti–correlation between β and r_g. If the parameter axes in Figure 25 are rotated so that they point along the major and minor axes of the ellipse then the two new parameters defined by the rotation are independent. One can then perform a single Karhunen–Loève eigenvalue compression for one of the new parameters, $\boldsymbol{\theta}' = \mathbf{R}\boldsymbol{\theta}$ (panel c). The parameter with the largest uncertainty should be chosen: this is the optimal–compression parameter. When rotating back to the original parameter axes after the compression (panel d), it is clear that the marginal uncertainties on both parameters have been unaffected so long as the original correlation has not changed.

The first step in applying GOMA, is to obtain the rotation matrix $\mathbf{R}$ by using the ellipse

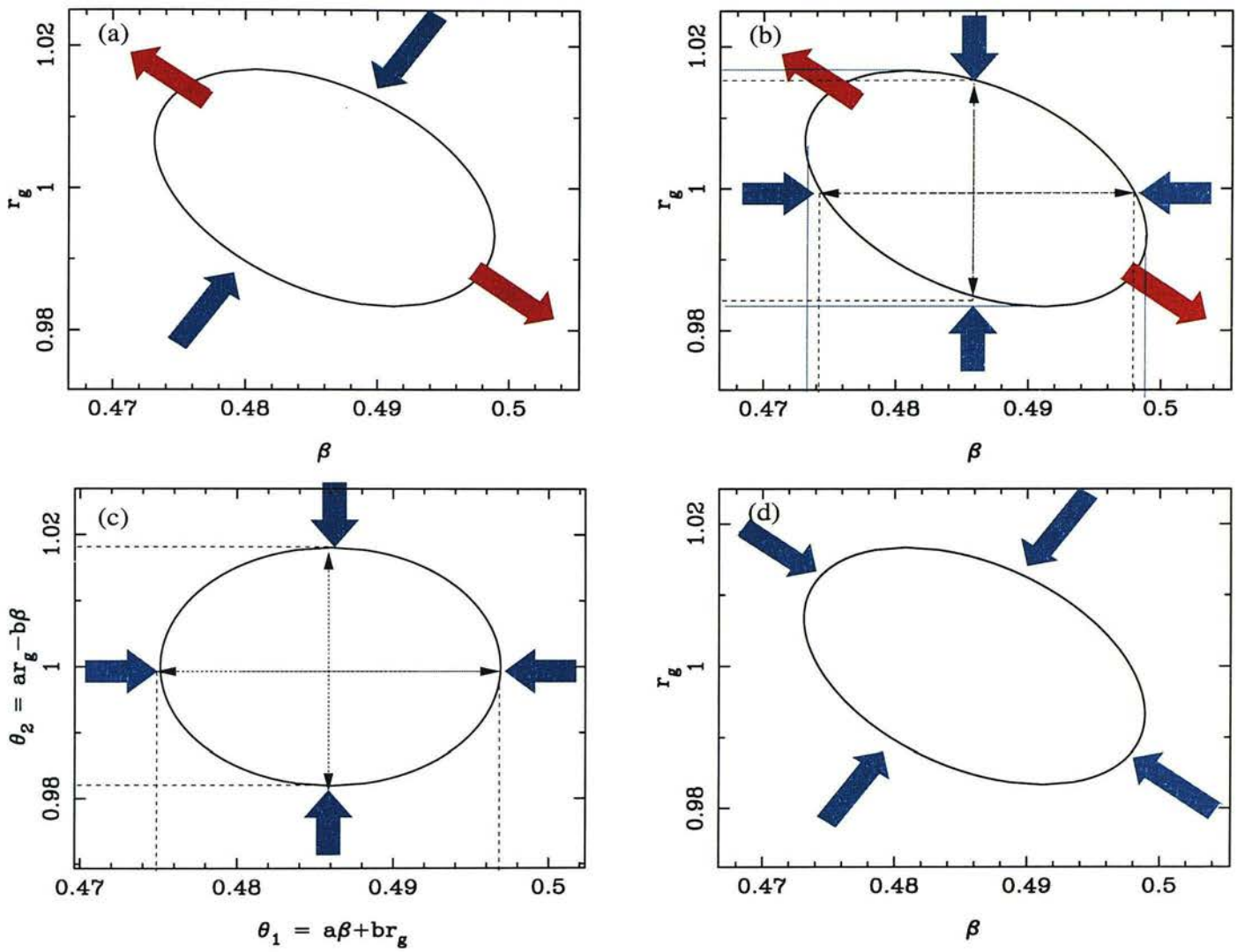

Figure 25: *A comparison of different schemes for data compression. Panel (a) shows a possible consequence of signal to noise compression. The compression resists changes in the area of the ellipse but may still lead to the stretching of the major axis. Panel (b) illustrates the Karhunen–Loève compression of Tegmark et al. The compression is optimal for constraining the* conditional *uncertainty on either parameter (marked by dashed lines), but since* β *and* r_g *are anti–correlated their* marginalised *uncertainties (marked by the solid lines) are greater. The compression is designed to pin down the conditional errors as indicated by the inward pointing arrows. However the ellipse can still stretch out along the major axis. Panel (c) shows the compromise offered by GOMA. The ellipse is rotated by defining two new independent parameters. The conditional uncertainty equals the marginal uncertainty for these parameters. Finally panel (d) shows that when the parameter vector is rotated back to the original parameters the compression has been optimal for the marginalised uncertainties.*

equation:

$$\boldsymbol{\theta}\mathbf{F}\boldsymbol{\theta}^t = \boldsymbol{\theta}\mathbf{R}\mathbf{F}'\mathbf{R}^t\boldsymbol{\theta}^t \tag{194}$$

where $\mathbf{F}'$ is the Fisher matrix of two new independent parameters $\boldsymbol{\theta}' = \mathbf{R}^t\boldsymbol{\theta}^t$. The diagonal matrix $\mathbf{F}'$ contains the eigenvalues of $\mathbf{F}$ in its diagonal elements. It then remains to choose the *smallest* eigenvalue $\mathbf{F}'_{ii}$ and perform the KL compression for the corresponding parameter $\boldsymbol{\theta}_{\rm KL} = \mathbf{R}_{1i}\boldsymbol{\theta}_1 + \mathbf{R}_{2i}\boldsymbol{\theta}_2$.

If the correlation of the original parameters changes sign after the compression, then the increased uncertainty in the second parameter will affect the marginal constraints. In this case it is necessary to apply the Karhunen Loève method to the second parameter as well (as is done in the figure). The choice of the new parameter space is still beneficial since for uncorrelated parameters the marginal uncertainty on a parameter is the same as its conditional error: the marginal errors are $\mathbf{F}_{ii}^{-1}$ which become $\frac{1}{\mathbf{F}_{ii}}$ for a diagonalised Fisher matrix.

To summarise, the steps of the compression are as follows;

1. first the Fisher matrix is constructed using the model covariance matrix for a set of fiducial parameters – the same model covaraince matrices that are used in the likelihood analysis can be used to construct the Fisher matrix using equation (220);

2. next the Fisher matrix is diagonalised to obtain the matrix $\mathbf{R}$;

3. the parameter $\mathbf{R_{1i}}\boldsymbol{\theta_1} + \mathbf{R_{2i}}\boldsymbol{\theta_2}$ is chosen as the parameter for which to optimally perform the compression, where the row i is chosen such that $\mathbf{F}'_{ii}$ is the smallest eigenvalue of the Fisher matrix;

4. finally the Karhunen Loève compression is performed for this parameter and if necessary for the second new parameter.

The steps above may be generalised to the case of many parameters. The matrix $\mathbf{C}$, $_i$ may be obtained either analytically or numerically, using the model power spectrum. The specifics of applying GOMA to the spherical harmonics models are described in Appendix C.2.

4.9 Practical Issues

4.9.1 Numerical roundoff errors and the choice of modes

In the cosmological model r_g by definition cannot exceed unity. This is not however included as a prior assumption since restricting the range of r_g biasses low any average values. In practice however it was found that the range of r_g that could be probed was limited by numerical problems in the model covariance matrix. As discussed in P04 problems arise when two modes with similar wavenumber are nearly degenerate. Numerical roundoff errors introduced by inversion or rotation processes can then make them completely degenerate leading to unphysical negative eigenvalues in the model covariance matrix. Furthermore for modes above a certain wavenumber, it was not even possible to perform the initial Cholesky decomposition in order to use the Karhunen–Loève compression on these modes, because the covariance matrix became non–positive definite. For this reason it was necessary to restrict the analysis to $k \leq 0.15\,h\mathrm{Mpc}^{-1}$: the upside being that the analysis is well within the limit of linear theory. Another consequence of the numerical roundoff problems is that negative eigenvalues were observed in the model covariance when r_g was increased above 1. The reason for this is not immediately apparent, but it is reasonable to conclude that the largest numerical errors are emerging from the parts of the model with an r_g dependence. Whereas with the 2dFGS analysis of P04 the major numerical problems arose due to a complicated survey mask, here the survey covers a uniform hemisphere. The problems for the 6dFGS are more likely to arise from the decomposition of the radial distribution into spherical Bessel functions. The velocity field is decomposed into spherical Bessel functions whose maximum wavelength fits into the larger range of the redshift survey. The velocity survey however, cuts off at the median depth of the redshift survey. For this reason, when the velocity field is decomposed into spherical Bessel functions the longest wavelength modes are all very similar. In order to ensure a positive definite covariance over the range of interest the following two steps were taken: a limited range of $r_g \leq 1$ was used, and the GOMA compression discussed in Section 4.8 was performed.

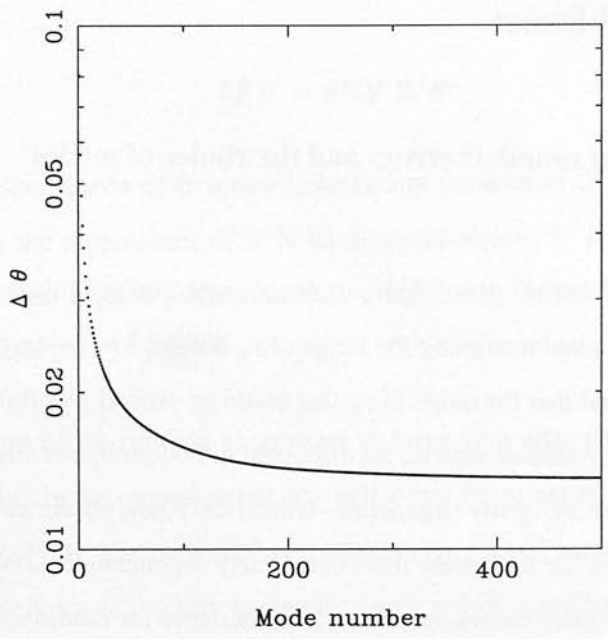

Figure 26: *Errors on the optimal compression parameter θ as a function of decreasing Karhunen–Loève eigenvalues. Errors are calculated using equation (235).*

4.9.2 The final compression

Limiting the wavenumber to $k_{max} = 0.15$, there are 2970 different sets of values (l, m, n), describing independent modes which fit the boundary conditions. Splitting the modes into their real and imaginary components multiplies this number by two. The combined data set for velocities and redshifts doubles this number again, giving a total of 11880 modes to work with. Clearly a brute force inversion of a $\sim 11000 \times 11000$ matrix is unfeasible but the number is greatly reduced by the azimuthal symmetry of a half (or full) sky survey. Symmetry means that modes of different m remain independent after convolution with the survey mask, therefore the data set can be defined as being the sum over different m values for each l, n without any loss of statistical information. This reduces the number of modes to a practical 195×4.

In order to reduce the numerical problems discussed above, it was then necessary to reduce this number further through the GOMA compression. Calculation of the Fisher information matrix from the model covariance matrix using:

$$\mathbf{F}_{ij} = \frac{1}{2}\mathbf{C}^{-1}\mathbf{C}_{,i}\,\mathbf{C}^{-1}\mathbf{C}_{,j} \tag{195}$$

(TTH, also see Appendix A.2) gives predictions of the attainable errors of $\Delta r_g = 0.017$ and

$\Delta\beta = 0.028$. The optimal parameter to compress for is found to be

$$\theta = 0.013 r_g + 0.99\beta. \tag{196}$$

Figure 26 shows the error bars expected on θ as calculated from the Karhunen–Loève eigenvalues (as explained in Appendix A.3). Only the first 65 modes were kept. This relatively conservative cutoff was chosen to eradicate the inversion problems already discussed. As can be seen from Figure 26 this resulted in the loss of very little information.

4.10 Results

4.10.1 Testing the formalism

Figures 27 and 28 show the results of the likelihood analysis applied to the simulations. For comparison the analysis was also performed on a 6dF–like survey with the same radial selection function but an all sky mask (see Figures 29 and 30). In all of the likelihood plots the fiducial values expected from the simulations ($r_g = 1.0, \beta = 0.54$) are plotted as a bold dot. The fiducial parameters are based on the assumption in constructing the catalogues that the galaxy field is unbiassed (so $r_g = 1$) and β is then obtained from the value of Ω_m input into the Hubble volume simulations on which the catalogues are based. A fundamental test of both the formalism and the software is whether the parameters constrained match the input parameters. Contours throughout correspond to the 1σ, 2σ and 3σ limits for two parameters, i.e. the 68.3%, 95.4% and 99.73% confidence regions. The dashed lines represent the 1σ boundaries of the marginalised errors for each parameter. Unlike Figure 25 the likelihood contours here are not perfect ellipses. This is partly because the parameters do not have a Gaussian distribution and partly because the contours are found to be mishaped by the compression. The result is that the marginal limits on each parameter lie inside the 2–parameter confidence regions.

Figures 31 and 32 show overall statistics for all 27 simulations, for the 6dFGS and all sky simulations respectively. The contours are a composite of the 27 individual likelihood functionals where the confidence regions have been rescaled to reflect what would be expected from an average realization of the survey. Also plotted are the positions of all 27 maxima.

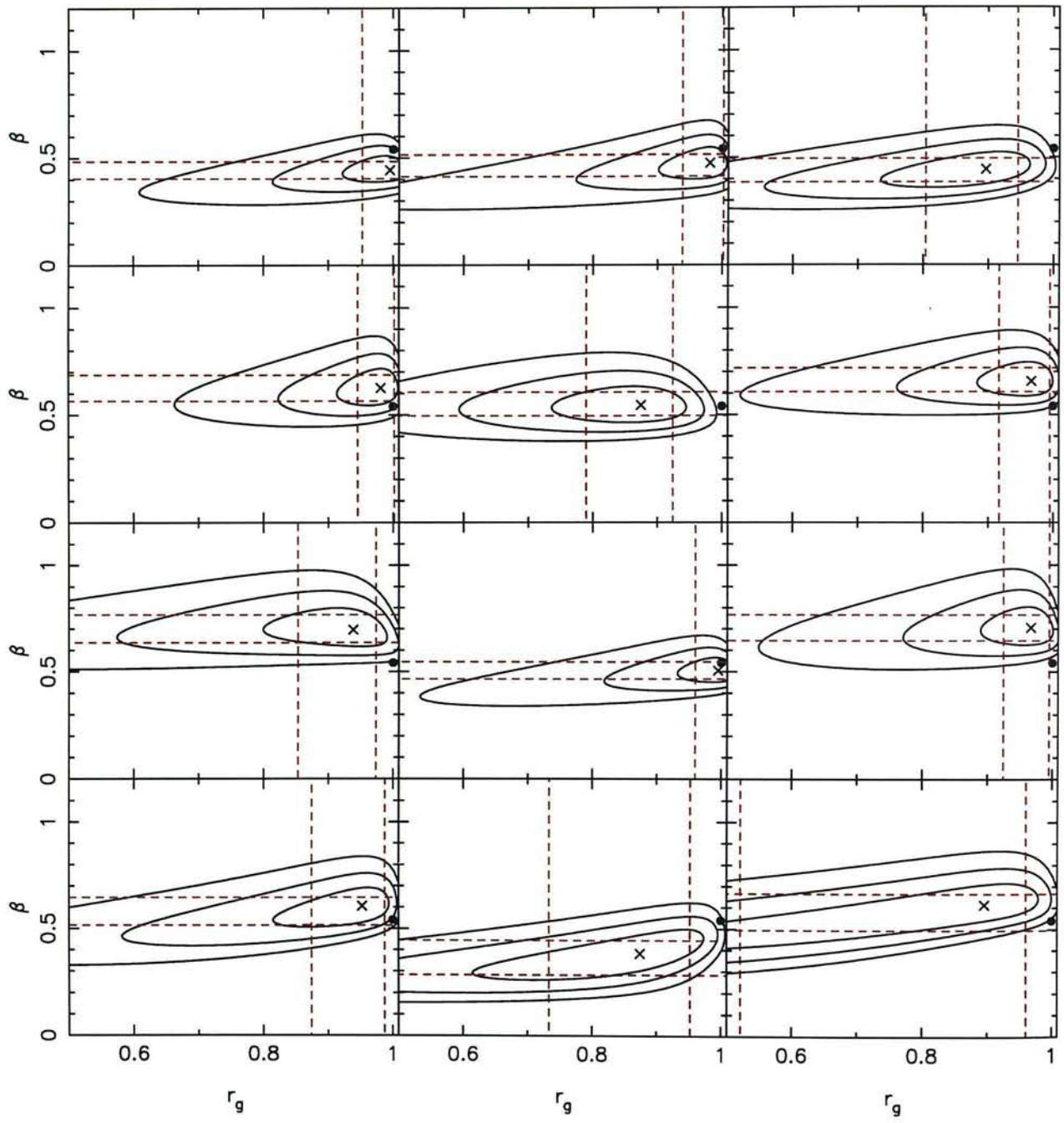

Figure 27: *Results of the analysis applied to simulations of the 6dFGS. 27 simulations were used in total, shown here and in Figure 28. The contours mark the* 1σ, 2σ *and* 3σ *limits for two parameters, i.e. the* 68.3%, 95.4% *and* 99.73% *confidence regions. The dashed lines mark the* 1σ *marginalized uncertainties for each parameter. The cross marks the maximum likelihood points from the analysis and the bold dot marks the fiducial value.*

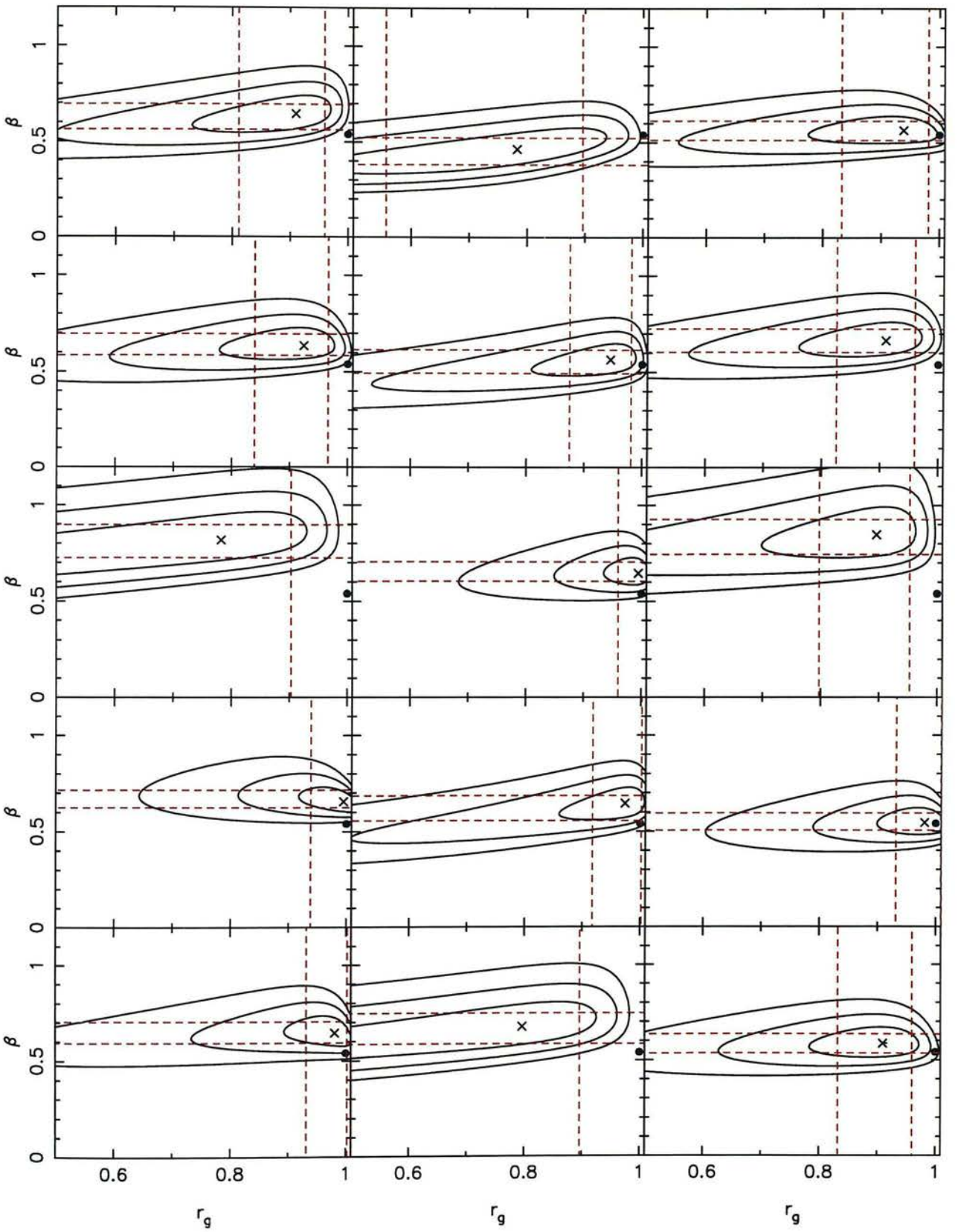

Figure 28: *6dFGS simulations continued.*

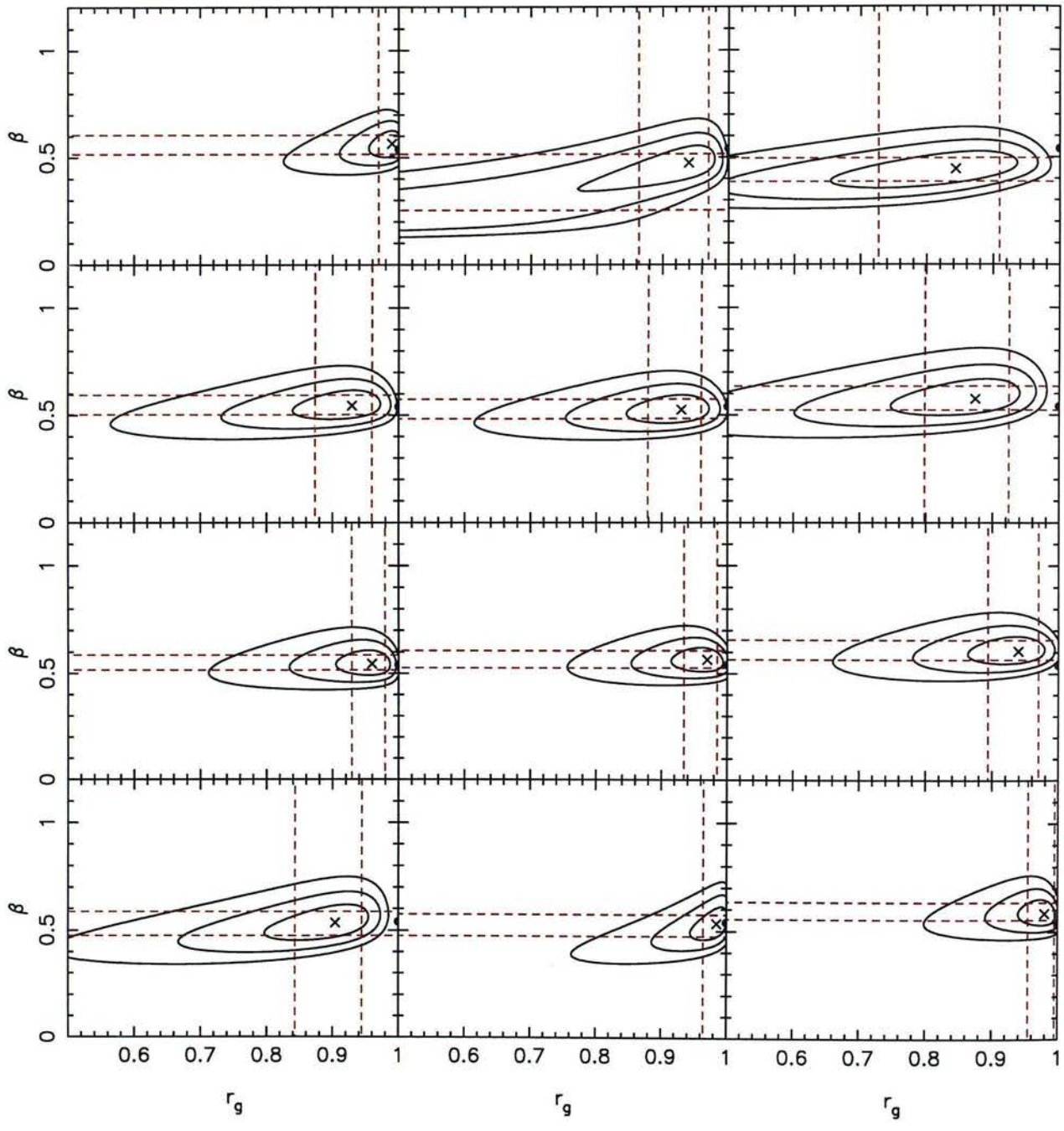

Figure 29: *The benefit of all sky surveys. The 27 simulations shown here and in Figure 30 are based on the 6dFGS radial selection functions but assuming an all sky mask. The parameter constraints are much improved by all sky coverage.*

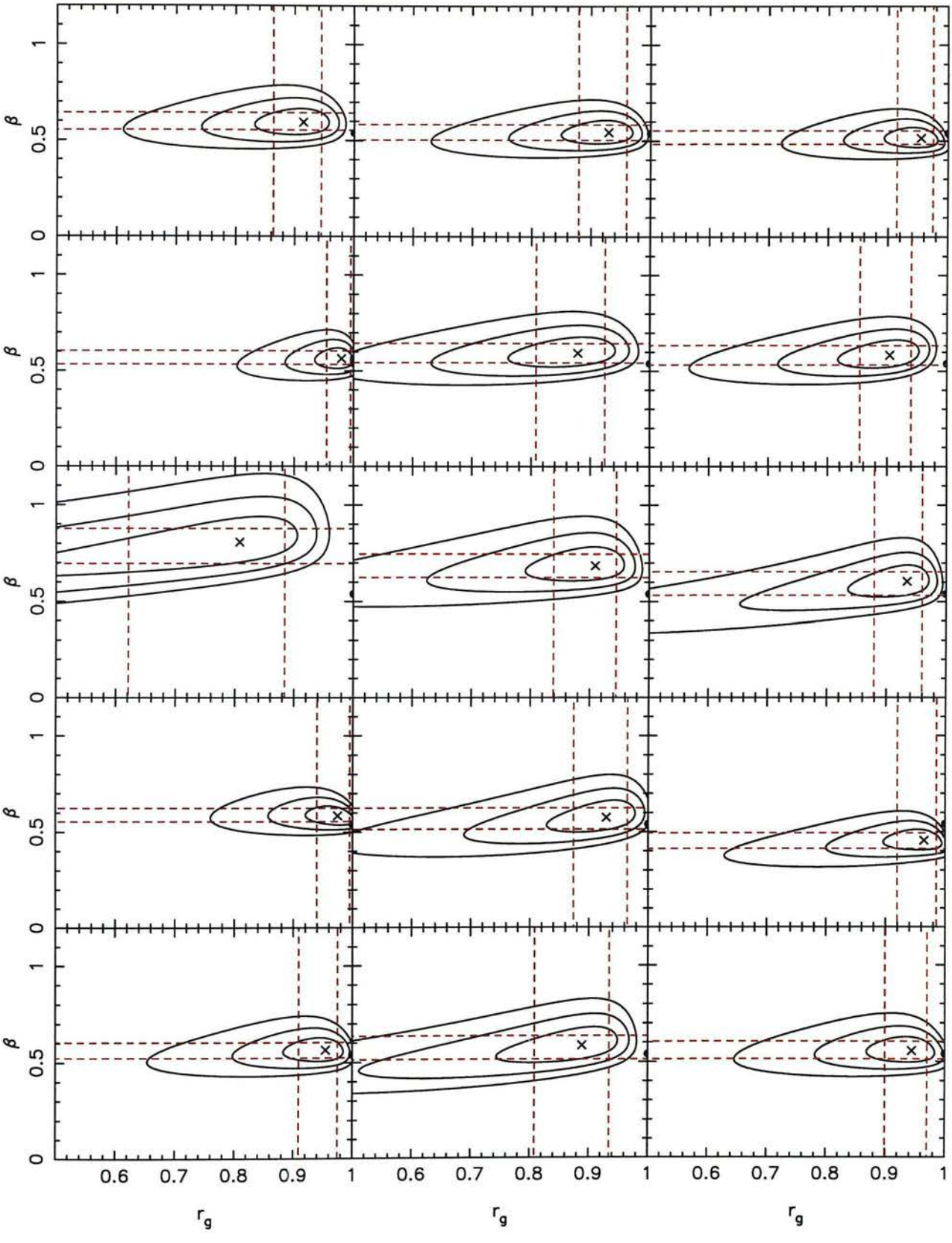

Figure 30: *All sky simulations continued.*

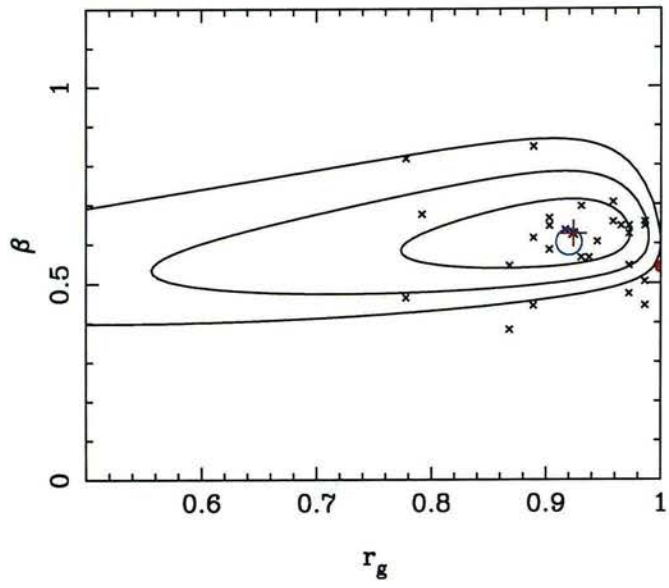

Figure 31: *Summary of the overall results of the 27 6dFGS half–sky simulations. The individual maximum likelihood points are shown (black crosses) as well as the average likelihood contours. The fiducial point is the filled red circle. Also shown are the average value of the ML points ($r_g = 0.920 \pm 0.061$, $\beta = 0.604 \pm 0.106$) , denoted by the open blue circle, and the ML point of the joint probability of the 27 realizations ($r_g = 0.924$, $\beta = 0.625$), marked by a red cross. The two points are in close agreement.*

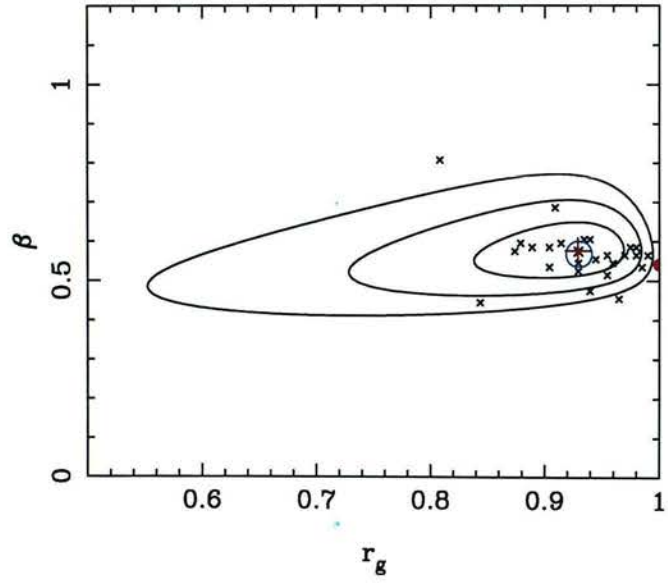

Figure 32: *Summary of the overall results of the 27 6dF–like all sky simulations. The contours are notably smaller than their half sky counterparts showing the advantage of a northern hemisphere followup to 6dFGS. The average value of the ML points is ($r_g = 0.930 \pm 0.043$, $\beta = 0.567 \pm 0.067$), and the ML point of the joint probability of the 27 realizations is ($r_g = 0.929$, $\beta = 0.575$). Again there is close agreement between the two. See Figure 31 for an explanation of the symbols.*

6dFGS simulations

Parameter	Joint estimate	Internal error	Averaged estimate	External error	Fiducial value
r_g	0.924	$^{+0.044}_{-0.198}$	0.920	± 0.061	1.0
β	0.625	$^{+0.085}_{-0.085}$	0.604	± 0.106	0.54

All Sky simulations

Parameter	Joint estimate	Internal error	Averaged estimate	External error	Fiducial value
r_g	0.929	$^{+0.035}_{-0.089}$	0.930	± 0.043	1.0
β	0.575	$^{+0.075}_{-0.075}$	0.567	± 0.067	0.54

Table 1: *Summary of the results for the compilation of simulations shown in Figures 31 and 32*

A comparison of these plots shows the benefit of performing an all–sky survey – the parameter constraints are significantly improved over the current 6dF design. Table 1 summarises the overall results for these compilations. As in Figures 31 and 32 there are two estimates quoted for each parameter. The joint estimate is the maximum likelihood point of all 27 likelihood surfaces multiplied together. This point is marked by a red cross in Figures 31 and 32. The error attached to this comes from the upper and lower boundaries of the 1σ confidence region for each parameter. This is a measure of the internal error, i.e. the error predicted by the model. The averaged likelihood is the mean position of the individual maximum likelihood points for each simulation. In the figures this is marked by a blue circle. The standard deviation of the individual maximum likelihood points is quoted in Table 1 as the external error. This error originates in the cosmic variance and shot noise of the simulations themselves. The similarity between internal and external error, and the near identical values of the averaged and joint estimates can be seen in Figures 31 and 32. This is an encouraging result as it implies that the model covariance and noise accurately describes the shape and width of the likelihood surface.

4.10.2 Tests for systematic problems

As already mentioned, the major systematic uncertainty to be overcome is the problem of numerical roundoff errors. In an analysis of a real survey one would need to carefully remove these problems through a selective choice of the modes. P04 achieve this by spacing the modes out by $\Delta l = 10$. The compression performed in this chapter is optimal in the sense that it discards data on the basis of its information content, however it does not necessarily discard

"problem modes" first and so it may be overcompressing the data in order to achieve stable results.

The first comment to be made about Figures 31 and 32 is that the method very accurately reproduces the fiducial value of β – this is to within $1-\sigma$ in both sets of simulations. One notable systematic however, is the slight offset between the predicted r_g and its fiducial value – formally this is a $2-\sigma$ offset but the likelihood contours are shaped such that the fiducial point lies at the edge of the 99.73% confidence region. There are three possible reasons for this discrepancy. One is that the simulations are not completely independent since they are drawn from the same Hubble volume simulation. An alternative possibility is that the model for small scale scattering is not quite precise enough. Given that the upper wavenumber of the analysis is well within the linear regime however, the effect of small scale scattering should be small. The most probable cause of the discrepancy is the numerical roundoff problems observed during data compression which have the effect of introducing negative eigenvalues into the model covariance matrix at $r_g > 1$. The problem of how to remove these problems in a systematic way, without throwing away information, will be of key interest in future analysis but its investigation is beyond the scope of the current chapter.

4.11 Summary

The formalism of HT has been extended to include the analysis of peculiar velocity fields. The bias between galaxies and dark matter has been modelled using the stochastic biassing correlation coefficient r_g introduced by Dekel and Lahav (1998). Both fields were decomposed into spherical harmonics and spherical Bessel functions in order to disentangle the effects of redshift distortions and mask convolution. Model covariance matrices were constructed for a joint data set of overdensity and velocity modes and a likelihood analysis was demonstrated by applying it to simulations of the 6dF Galaxy Survey.

The method requires that one first perform a compression of the data using GOMA. The optimal parameter to compress for in this case is found to be very nearly equal to β. It was found that in the absence of the need to reject certain modes the analysis should be able to constrain β to $\sim 5\%$ and r_g to $\sim 2\%$. This was calculated using equation (195) by substituting

in the model covariance matrices and noise derived in Sections 4.3 and 4.5. Incidentally this provides an independent verifcation of the predictions made in Chapter 3 about the parameter constraints achievable from 6dFGS. Comparison with Figure 21 shows that these values are in rough agreement.

The main test of the analysis is whether it returns the parameter values input into the simulations. In this respect the method is successful in returning confidence regions which encompass the fiducial value. The marginalised error bars on each estimate are typically of order $\sim 10\%$ for β and $\sim 5\%$ for r_g although this varies significantly from one simulation to the next. These constraints are of the order predicted in Chapter 3 for the 6dFGS at the limiting wavenumber of $k = 0.15h\mathrm{Mpc}^{-1}$. However it is recognised that numerical problems in inverting and rotating matrices has restricted the amount of modes available for use in the analysis and so reduced the useable information content of the data.

When the 6dFGS is complete and the time comes to apply this analysis to the real data, it will be necessary to perform a similar analysis to that shown above, using the exact window function and selection function of the survey, and performing a more careful treatment of numerical inversion problems. For now the analysis shown above, suffices to demonstrate the consistency of the formalism with the simulations, and to demonstrate the parameter constraint capabilities of the method applied to the 6dFGS.

5 Conclusions

5.1 Summary

This thesis has concentrated on the statistical information on cosmological parameters contained within galaxy redshift and radial peculiar velocity surveys. The particular example of the 6dF Galaxy Survey has been used, since it constitutes the first large combined survey of redshifts and $D_n - \sigma$ distances. The research presented has been divided into two distinct parts. Chapter 3 concerned predictions of the information content of the 6dFGS survey. Chapter 4 developed a maximum likelihood approach for extracting the information from redshifts and peculiar velocities, using spherical basis functions to describe the fields.

Chapter 3 involved use of the Fisher information matrix to predict the uncertainites within which galaxy surveys can predict parameters of a model power spectrum. Initially redshift surveys were discussed. An expression for the shot noise expected from the 6dF redshift survey was obtained through modelling the expected galaxy number counts. The geometry of the survey design was parameterised with a sky fraction, a median depth and a completeness parameter. The optimal design was then obtained by finding the combination of these parameters which minimizes the fractional uncertainty on the power spectrum amplitude, as predicted by the Fisher matrix. The optimal design was found to have as wide an angular coverage as possible at the expense of going deep. The optimal median depth was found to be $r_m = 230h^{-1}\mathrm{Mpc}$ which is different to the expected value of the survey, but it was found that it made little difference in practice. Using these optimal design parameters, marginalised error bars for a few different combinations of parameters were found. The results show that a combination of the galaxy power spectrum amplitude, the distortion parameter and the shape parameter can all be constrained to a few percent. However the effect of including a fourth free parameter – either the mass–galaxy correlation coefficient r_g or the baryonic density parameter, is that the extra parameter cannot be constrained and its inclusion induces extra uncertainty in the constraint of the first three parameters. Analytical results were also presented for the uncertainty in the distortion parameter in the limit of negligible shot noise. It was found that to achieve a 1 % constraint on β requires a survey some four times the volume of the 6dFGS.

Chapter 3 then saw the analysis applied to velocity surveys. A similar optimisation of the survey design was performed, this time taking into account the distance indicator scatter in the $D_n - \sigma$ relation. The optimal velocity survey design was found to have as accurate distance estimates as possible at the expense of depth. The velocity survey alone will be able to jointly constrain parameters A_v and Γ to $\sim 25\%$ and it will be able to obtain a conditional constraint of A_v to $\sim 6\%$.

The final subject of Chapter 3 was the combination of velcoities and redshifts. The Fisher matrix for the combination of velocity and density fields was derived. This was then used to obtain predictions of the parameter estimation capabilities of a joint analysis. The greatest improvement was found to be in the constraint of the parameter r_g. The joint constraint of r_g with β becomes possible, and this was shown to be due to the breaking of the degeneracy between the two parameters inherent in the redshift–only analysis. Plots of the expected likelihood contours showed that the inclusion of velocities makes these parameters independent. A joint analysis of redshifts and peculiar velocities from the 6dFGS was shown to be capable of parameter constraints for A_g, Γ, β and r_g all at the level of a few percent.

Chapter 4 involved the construction of a maximum likelihood method to constrain r_g and β using a combination of velocities and redshifts. The spherical harmonics analysis of Heavens and Taylor (1995) was extended to include radial velocities. The method involved expanding both density and velocity fields in spherical harmonics and spherical Bessel functions. The data set was then defined as the set of Fourier modes of this expansion. The spherical basis allowed a separation of angular and radial effects. Convolution of modes with the window function of the survey destroys their independence so that full construction of their covariance matrix was necessary. Careful treatment was given to redshift space distortions. In the spherical harmonics formalism it is possible to perform an exact treatment of this effect to linear order improving on the plane wave approximation of Kaiser (1986).

A model for the spherical modes of the velocity field was constructed in Chapter 4. The modes were chosen to be those of the radial gradient of the radial velocity field since these modes are correlated with the density field modes. To obtain these radial gradient modes from the observational peculiar velocities it was necessary to integrate by parts. The shot noise on the peculiar velocity modes was also found. The inclusion of $D_n - \sigma$ errors can be achieved by

adding the resultant velocity variance in quadrature to the intrinsic scatter in peculiar velocities. The models presented in Chapter 4 were used to construct the full covariance matrix of density and velocity modes for use in a likelihood analysis.

Chapter 4 then discussed practical problems in performing a likelihood analysis. Data compression techniques were compared and the Generalised Optimal Mode Analysis technique was applied to the spherical harmonics modes. The chapter also discussed the numerical problems stemming from matrix inversion and rotation for nearly degenrate modes.

Finally Chapter 4 showed the results of the spherical harmonics analysis applied to mock 6dFGS catalogues constructed from the Hubble Volume simulations. The analysis was found to recover the value of β input into the simulations to within 1σ and r_g to within 2σ. The internal and external errors were in close agreement confirming the validity of the model covariance and noise as a description of the likelihood surface. Numerical problems were also found to require the discarding of some information, leading to individual error bars being slightly larger than predicted in Chapter 3.

5.2 Future Applications

The spherical harmonics software is now fully prepared to be used on the 6dFGS data sets when the velocity survey is complete. The programs can generally be used on a data set given an input sky mask and an accurate selection function. The issue of numerical roundoff problems will need to be properly addressed however. A careful treatment of the problems of degenerate modes will involve painstaking preparation of the input data set, selectively chosen to avoid these problems, as was done by Percival et al. (2004). The length of time taken to construct the model covariance matrix – in particular the construction of the velocity component and the small scale scattering matrix – means that trial and error application of the procedure on real data will be long and hard. The exact treatment of linear redshift distortions however, makes this the most physically accurate method of constraining β, and with r_g left as a free parameter, the constriant of β is also free of assumptions about the biassing relationship.

Completion of the 6dF Galaxy Survey is expected in the middle of 2005.

References

Aaronson, M. et al.: 1982, *ApJS* **50**, 241

Abazajian, K.: 2003, *AJ* **126**, 2081

Abramowitz, M. and Stegun, I. A.: 1972

Alcock, C. and Paczynski, B.: 1979, *Nature* **281**, 358

Ballinger, W. E., Taylor, A. N., Heavens, A. F., and Tadros, H.: 2000

Bardeen, J. M., Bond, J. R., Kaiser, N., and Szalay, A. S.: 1986, *ApJ* **304**, 15

Benoist, C., Maurogordato, S., da Costa, L. N., Cappi, A., and Schaeffer, R.: 1996, *ApJ* **472**, 452

Burles, S.: 2002, *Planetary and Space Science* **50**, 1245

Carroll, S. M., Press, W. H., and Turner, E. L.: 1992, *Annual Review of Astronomy and Astrophysics* **30**, 499

Cole, S. et al.: 2001, *MNRAS* **326**, 255

Coles, P. and Jones, B.: 1991, *MNRAS* **248**, 1

Colless, M. et al.: 1999, in *Looking Deep in the Southern Sky*, p. 9

Colless, M., Jones, H., Campbell, L., Burkey, D., Taylor, A. N., and Saunders, W.: 2004, *in press*

Colless, M., Saglia, R. P., Burstein, D., Davies, R. L., McMahan, R. K., and Wegner, G.: 2000, in *ASP Conf. Ser. 201: Cosmic Flows Workshop*, p. 54

Courteau, S. and Dekel, A.: 2001, in *ASP Conf. Ser. 245: Astrophysical Ages and Times Scales*, p. 584

da Costa, L. N. et al.: 1988, *ApJ* **327**, 544

da Costa, L. N. et al.: 1994, *ApJl* **424**, L1

de Lapparent, V., Geller, M. J., and Huchra, J. P.: 1986, *ApJl* **302**, L1

Dekel, A., Bertschinger, E., Yahil, A., Strauss, M. A., Davis, M., and Huchra, J. P.: 1993, *ApJ* **412**, 1

Dekel, A. and Lahav , O.: 1999, *ApJ* **520**, 24

Dekel, A. and Rees, M. J.: 1994, *ApJl* **422**, L1

Djorgovski, S. and Davis, M.: 1987, *ApJ* **313**, 59

Dressler, A.: 1980, *ApJ* **236**, 351

Dressler, A. et al.: 1987a, *ApJl* **313**, L37

Dressler, A. et al.: 1987b, *ApJ* **313**, 42

Dressler, A. and Faber, S. M.: 1990, *ApJl* **354**, L45

Efstathiou, G. et al.: 2002, *MNRAS* **330**, L29

Eisenstein, D. J. and Hu , W.: 1998, *ApJ* **496**, 605

Erdogdu, P. et al.: 2004, *MNRAS* **352**, 939

Faber, S. M. and Burstein, D.: 1988, in *Large-Scale Motions in the Universe: A Vatican study Week*, pp 115–167

Faber, S. M. and Jackson, R. E.: 1976, *ApJ* **204**, 668

Feldman, H. A., Kaiser, N., and Peacock, J. A.: 1994, *ApJ* **426**, 23

Fisher, K. B., Davis, M., Strauss, M. A., Yahil, A., and Huchra, J. P.: 1994a, *MNRAS* **267**, 927

Fisher, K. B., Huchra, J. P., Strauss, M. A., Davis, M., Yahil, A., and Schlegel, D.: 1995a, *ApJS* **100**, 69

Fisher, K. B., Lahav, O., Hoffman, Y., Lynden-Bell, D., and Zaroubi, S.: 1995b, *MNRAS* **272**, 885

Fisher, K. B., Scharf, C. A., and Lahav, O.: 1994b, *MNRAS* **266**, 219

Fisher, K. B., Scharf, C. A., and Lahav, O.: 1994c, *MNRAS* **266**, 219

Freedman, W. L. et al.: 2001, *Apj* **553**, 47

Freedman, W. L., Madore, B. F., and Kennicutt, R. C.: 1997, in *The Extragalactic Distance Scale*, pp 171–185

Frith, W. J., Busswell, G. S., Fong, R., Metcalfe, N., and Shanks, T.: 2003, *MNRAS* **345**, 1049

Geller, M. J. et al.: 1997, *AJ* **114**, 2205

Geller, M. J. and Huchra, J. P.: 1989, *Science* **246**, 897

Gorski, K. M., Davis, M., Strauss, M. A., White, S. D. M., and Yahil, A.: 1989, *ApJ* **344**, 1

Gregg , M. D.: 1995, *AJ* **110**, 1052

Gregory, S. A. and Thompson, L. A.: 1978, *Nature* **274**, 450

Hamilton, A. J. S.: 1998, in *ASSL Vol. 231: The Evolving Universe*, p. 185

Hawkins, E. et al.: 2003, *MNRAS* **346**, 78

Heath, D. J.: 1977, *MNRAS* **179**, 351

Heavens, A. F. and Taylor , A. N.: 1995, *MNRAS* **275**, 483

Heavens, A. F. and Taylor , A. N.: 1997, *MNRAS* **290**, 456

Hoessel, J. G.: 1980, *ApJ* **241**, 493

Hubble, E.: 1929, *Proceedings of the National Academy of Science* **15**, 168

Huchra, J., Davis, M., Latham, D., and Tonry, J.: 1983, *ApJS* **52**, 89

Hudson, M. J., Colless, M., Dressler, A., and Giovanelli, R.: 2000, in *ASP Conf. Ser. 201: Cosmic Flows Workshop*, p. 159

Jarrett, T. H., Chester, T., Cutri, R., Schneider, S., Skrutskie, M., and Huchra, J. P.: 2000, *AJ* **119**, 2498

Jorgensen, I., Franx, M., and Kjaergaard, P.: 1993, *ApJ* **411**, 34

Kaiser, N.: 1984, *ApJl* **284**, L9

Kaiser, N.: 1987a, *MNRAS* **227**, 1

Kaiser, N.: 1987b, *MNRAS* **227**, 1

Kendall, M. G. and Stuart , A.: 1969, *The Advanced Theory of Statistics*, Vol. 2, Griffin, London

Kochanek, C. S. et al.: 2001, *ApJ* **560**, 566

Kolatt, T. and Dekel, A.: 1994, *ApJ* **428**, 35

Lahav, O.: 2002

Lahav, O. et al.: 2002, *MNRAS* **333**, 961

Lahav, O., Lynden-Bell, D., and Rowan-Robinson, M.: 1988, *MNRAS* **234**, 677

Lahav, O., Rees, M. J., Lilje, P. B., and Primack, J. R.: 1991, *MNRAS* **251**, 128

Lauer, T. R. and Postman, M.: 1994, *ApJ* **425**, 418

Lawrence, C. R. and Lange, A. E.: 1997, *Bulletin of the American Astronomical Society* **29**, 1273

Loveday, J., Efstathiou, G., Peterson, B. A., and Maddox, S. J.: 1992, *ApJl* **400**, L43

Lynden-Bell, D. et al.: 1988, *ApJ* **326**, 19

Madgwick, D. S. et al.: 2003, *MNRAS* **344**, 847

Mather, J. C., Fixsen, D. J., Shafer, R. A., Mosier, C., and Wilkinson, D. T.: 1999, *Apj* **512**, 511

Mathewson, D. S., Ford, V. L., and Buchhorn, M.: 1992, *ApJl* **389**, L5

Meiksin, A. and Davis, M.: 1986, *Aj* **91**, 191

Norberg, P. et al.: 2002, *MNRAS* **336**, 907

Parker, Q. A., Colless, M., and Mamon, G.: 1997, in *ASSL Vol. 212: Wide-field spectroscopy*, p. 303

Peacock, J. A. and Dodds , S. J.: 1994, *MNRAS* **267**, 1020

Peacock, J. A. et al.: 2001, in *Deep Fields*, p. 221

Peebles , P. J. E.: 1980, *The Large Scale Structure of the Universe*, Princeton University Press, Princeton N.J.

Penzias, A. A. and Wilson, R. W.: 1965, *ApJ* **142**, 419

Percival, W. J. et al.: 2001, *MNRAS* **327**, 1297

Percival, W. J. et al.: 2002, *MNRAS* **337**, 1068

Percival, W. J. et al.: 2004

Perlmutter, S., Turner, M. S., and White, M.: 1999, *Physical Review Letters* **83**, 670

Riess, A. G.: 2000, in *ASP Conf. Ser. 201: Cosmic Flows Workshop*, p. 80

Rubin, V. C. and Ford, W. K.: 1970, in *IAU Symp. 38: The Spiral Structure of our Galaxy*, p. 61

Rubin, V. C., Roberts, M. S., Graham, J. A., Ford, W. K., and Thonnard, N.: 1976, *AJ* **81**, 687

Rybicki, G. B. and Press, W. H.: 1992, *ApJ* **398**, 169

Sandage, A. and Tammann, G. A.: 1981, in *Carnegie Inst. of Washington, Publ. 635; Vol. 0; Page 0*, p. 0

Santiago, B. X., Strauss, M. A., Lahav, O., Davis, M., Dressler, A., and Huchra, J. P.: 1995, *ApJ* **446**, 457

Saunders, W. et al.: 1998, in *Wide Field Surveys in Cosmology, 14th IAP meeting held May 26-30, 1998, Paris. Publisher: Editions Frontieres. ISBN: 2-8 6332-241-9, p. 71.*, p. 71

Saunders, W., Frenk, C., Rowan-Robinson, M., Lawrence, A., and Efstathiou, G.: 1991, *Nature* **349**, 32

Schmoldt, I. M. et al.: 1999, *AJ* **118**, 1146

Seljak, U. and Bertschinger, E.: 1994, *ApJ* **427**, 523

Shectman, S. A. et al.: 1996, *ApJ* **470**, 172

Silberman, L., Dekel, A., Eldar, A., and Zehavi, I.: 2001, *ApJ* **557**, 102

Silk, J.: 1968, *ApJ* **151**, 459

Smoot, G. F.: 1999, in *AIP Conf. Proc. 476: 3K cosmology*, p. 1

Spergel, D. N. et al.: 2003, *ApJ* **148**, 175

Strauss, M. A., Cen, R., and Ostriker, J. P.: 1993, *ApJ* **408**, 389

Strauss, M. A., Cen, R., Ostriker, J. P., Lauer, T. R., and Postman, M.: 1995, *ApJ* **444**, 507

Strauss, M. A. et al.: 1992, *ApJ* **397**, 395

Sugiyama, N.: 1995, *ApJS* **100**, 281

Sutherland, W. et al.: 1999, *MNRAS* **308**, 289

Tadros, H. et al.: 1999, *MNRAS* **305**, 527

Taylor, A. N., Ballinger, W. E., Heavens, A. F., and Tadros, H.: 2001a, *MNRAS* **327**, 689

Taylor, A. N., Ballinger, W. E., Heavens, A. F., and Tadros, H.: 2001b, *MNRAS* **327**, 689

Taylor, A. N. and Watts, P. I. R.: 2001, *MNRAS* **328**, 1027

Tegmark, M.: 2003a, *astro-ph/0310723*

Tegmark, M.: 2003b, *astro-ph/0310725*

Tegmark, M., Taylor, A. N., and Heavens, A. F.: 1997, *ApJ* **480**, 22

Tonry, J. L., Dressler, A., Blakeslee, J. P., and Ajhar, E. A.: 2000, in *ASP Conf. Ser. 201: Cosmic Flows Workshop*, p. 70

Tully, R. B. and Fisher, J. R.: 1977, *A&A* **54**, 661

van Leeuwen, F.: 1996, *Space Science Reviews* **81**, 201

Verde, L. et al.: 2002, *MNRAS* **335**, 432

Vettolani, G. et al.: 1997, *A&A* **325**, 954

Wakamatsu, K., Colless, M., Jarrett, T., Parker, Q., Saunders, W., and Watson, F.: 2003, in *ASP Conf. Ser. 289: The Proceedings of the IAU 8th Asian-Pacific Regional Meeting, Volume I*, pp 97–104

Wegner, G., da Costa, L. N., Alonso, M. V., Bernardi, M., Wilmer, C. N. A., Pellegrini, P. S., Rité, C., and Maia, M.: 2000, in *ASP Conf. Ser. 201: Cosmic Flows Workshop*, p. 62

Willick, J. A.: 1990, *ApJl* **351**, L5

Willick, J. A.: 1999, in *Formation of Structure in the Universe*, p. 213

Willmer, C. N. A., da Costa, L. N., and Pellegrini, P. S.: 1998, *AJ* **115**, 869

Yahil, A., Walker, D., and Rowan-Robinson, M.: 1986, *ApJl* **301**, L1

Zaroubi, S., Hoffman, Y., Fisher, K. B., and Lahav, O.: 1995, *ApJ* **449**, 446

Zehavi, I. et al.: 2002, *ApJ* **571**, 172

Zeldovich, Y. B.: 1967, *JETP Lett.* **6**, 316

Zwicky, F.: 1937, *ApJ* **86**, 217

A Useful results for the Fisher matrix

A.1 The Cramér–Rao inequality

Let $L(\mathbf{x}|\boldsymbol{\theta})$ be the probability of obtaining data set $\mathbf{x}$ given model parameters $\boldsymbol{\theta}$. This probability is normalised to one:

$$\int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} \prod_{j=1}^{N} L(\mathbf{x}|\boldsymbol{\theta}) dx_j = 1 \tag{197}$$

where N is the number of data points. Differentiating with respect to one of the model parameters, and writing $\frac{\partial L}{\partial \theta_i} = \frac{\partial \log L}{\partial \theta_i} L$, gives:

$$\int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} \prod_{j=1}^{N} \frac{\partial \log L}{\partial \theta_i} L dx_j = 0 \tag{198}$$

which says that the expectation value of the log likelihood is zero:

$$\left\langle \frac{\partial \log L}{\partial \theta_i} \right\rangle = 0. \tag{199}$$

Differentiating a second time gives:

$$\int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} \prod_{j=1}^{N} dx_j \left(\frac{\partial^2 \log L}{\partial \theta_i^2} + \frac{\partial \log L}{\partial \theta_i} \frac{\partial \log L}{\partial \theta_i} \right) L = 0 \tag{200}$$

or

$$\left\langle \frac{\partial^2 \log L}{\partial \theta_i^2} \right\rangle = \left\langle -\left(\frac{\partial \log L}{\partial \theta_i} \right)^2 \right\rangle. \tag{201}$$

Let $\theta_{\rm est}$ be an unbiassed estimate of θ_i, then:

$$\int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} \prod_{j=1}^{N} \theta_{\rm est} L dx_j = \theta_i. \tag{202}$$

Differentiating this with respect to θ_i gives:

$$\int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} \prod_{j=1}^{N} \theta_{\rm est} \frac{\partial \log L}{\partial \theta_i} L dx_j = 1. \tag{203}$$

Using equation (198) this can be written:

$$\int_{-\infty}^{\infty} \ldots \int_{-\infty}^{\infty} \prod_{j=1}^{N} (\theta_{\rm est} - \theta_i) \frac{\partial \log L}{\partial \theta_i} L dx_j = 1. \tag{204}$$

At this stage it is useful to refer to the Cauchy-Schwarz inequality which states that for any functions f and g:

$$\int f(x)g(x) \leq \sqrt{\int f(x)^2 dx \int g(x)^2 dx}. \tag{205}$$

Defining f and g as:

$$\begin{aligned} f(x) &= (\theta_{\rm est} - \theta_i)\sqrt{L} \\ g(x) &= \frac{\partial \log L}{\partial \theta_i}\sqrt{L} \end{aligned} \tag{206}$$

equations (204) and (205) give:

$$1 = \int f(x)g(x)d\mathbf{x} \leq \sqrt{\int f(x)^2 d\mathbf{x} \int g(x)^2 d\mathbf{x}} \tag{207}$$

or in other words:

$$1 \leq \sqrt{\left(\int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} \prod_{j=1}^{N} dx_j (\theta_{\rm est} - \theta_i)^2 L\right)} \times \sqrt{\left(\int_{-\infty}^{\infty} ... \int_{-\infty}^{\infty} \prod_{j=1}^{N} dx_j \left(\frac{\partial \log L}{\partial \theta_i}\right)^2 L\right)}. \tag{208}$$

Rearranging to get the final result:

$$\langle (\theta_{\rm est} - \theta_i)^2 \rangle^{1/2} \geq \left\langle \left(\frac{\partial \log L}{\partial \theta_i}\right)^2 \right\rangle^{1/2} \tag{209}$$

where the left hand side is just the standard deviation of estimates of θ_i. Finally equation (201) can be used to give the minimum standard deviation in terms of the Fisher matrix as:

$$\Delta\theta_i \geq \mathbf{F}_{ii}^{-1/2} \tag{210}$$

using the definition of the Fisher matrix in equation (96). This result generalises to

$$\Delta\theta_i \geq \sqrt{(\mathbf{F^{-1}})_{ii}} \tag{211}$$

for the case of degenerate parameters. This is the Cramér–Rao inequality.

A.2 The Fisher matrix from the data covariance matrix

The proof outlined in this appendix is taken from Tegmark et al. (1997). The algebra is somewhat cumbersome but nevertheless provides an expression of fundamental importance to the Fisher matrix analysis in Chapter 3.

The Fisher matrix is defined as :

$$\mathbf{F}_{ij} = \left\langle \frac{\partial^2(-\ln L)}{\partial\theta_i \partial\theta_j} \right\rangle. \tag{212}$$

Setting $L(\boldsymbol{\theta}|\mathbf{x}) = L(\mathbf{x}|\boldsymbol{\theta})$ using Bayes theorem with equal priors, the log likelihood, written $\mathcal{L}$ is:

$$2\mathcal{L} = \ln\det \mathbf{C} + (\mathbf{x}-\boldsymbol{\mu})\mathbf{C}^{-1}(\mathbf{x}-\boldsymbol{\mu})^t \tag{213}$$

for a Gaussian distribution. The mean $\boldsymbol{\mu}$ and covariance $\mathbf{C} = \langle(\mathbf{x}-\boldsymbol{\mu})(\mathbf{x}-\boldsymbol{\mu})^t\rangle$ both may depend on the set of model parameters $\boldsymbol{\theta}$. Defining the data matrix as

$$\mathbf{D} \equiv (\mathbf{x}-\boldsymbol{\mu})(\mathbf{x}-\boldsymbol{\mu})^t \tag{214}$$

and using the matrix identity $\ln\det \mathbf{C} = \mathrm{Tr}\ln \mathbf{C}$ equation (213) becomes:

$$2\mathcal{L} = \mathrm{Tr}[\ln \mathbf{C} + \mathbf{C}^{-1}\mathbf{D}]. \tag{215}$$

It is convenient to define a comma notation for derivatives such that:

$$\mathbf{C}_{,i} \equiv \frac{\partial}{\partial\theta_i}\mathbf{C}. \tag{216}$$

Using the matrix identities $(\mathbf{C}^{-1})_{,i} = -\mathbf{C}^{-1}\mathbf{C}_{,i}\mathbf{C}^{-1}$ and $(\ln \mathbf{C})_{,i} = \mathbf{C}^{-1}\mathbf{C}_{,i}$, equation (215) can be differentiated:

$$2\mathcal{L}_{,i} = \mathrm{Tr}[\mathbf{C}^{-1}\mathbf{C}_{,i} - \mathbf{C}^{-1}\mathbf{C}_{,i}\mathbf{C}^{-1}\mathbf{D} + \mathbf{C}^{-1}\mathbf{D}_{,i}]. \tag{217}$$

When evaluated at the true parameter values $\langle\mathbf{x}\rangle = \boldsymbol{\mu}$ and $\langle\mathbf{x}\mathbf{x}^t\rangle = \mathbf{C} + \boldsymbol{\mu}\boldsymbol{\mu}^t$ which gives:

$$\begin{aligned} \langle\mathbf{D}\rangle &= \mathbf{C}, \\ \langle\mathbf{D}_{,i}\rangle &= 0, \\ \langle\mathbf{D}_{,ij}\rangle &= \boldsymbol{\mu}_{,i}\boldsymbol{\mu}^t_{,j} + \boldsymbol{\mu}_{,j}\boldsymbol{\mu}^t_{,i}. \end{aligned} \tag{218}$$

Differentiating equation (217) a second time, and using the chain rule gives:

$$\begin{aligned} 2\mathcal{L}_{,ij} = \mathrm{Tr}[&-\mathbf{C}^{-1}\mathbf{C}_{,i}\mathbf{C}^{-1}\mathbf{C}_{,j} + \mathbf{C}^{-1}\mathbf{C}_{,ij} + \mathbf{C}^{-1}(\mathbf{C}_{,i}\mathbf{C}^{-1}\mathbf{C}_{,j} + \mathbf{C}_{,j}\mathbf{C}^{-1}\mathbf{C}_{,i})\mathbf{C}^{-1}\mathbf{D} \\ &-\mathbf{C}^{-1}(\mathbf{C}_{,i}\mathbf{C}^{-1}\mathbf{D}_{,j} + \mathbf{C}_{,j}\mathbf{C}^{-1}\mathbf{D}_{,i}) - \mathbf{C}^{-1}\mathbf{C}_{,ij}\mathbf{C}^{-1}\mathbf{D} + \mathbf{C}^{-1}\mathbf{D}_{,ij}]. \end{aligned} \tag{219}$$

Substituting this into equation (212), using the results of equation (218) and using the trace identity that $\mathrm{Tr}[AB] = \mathrm{Tr}[BA]$, the Fisher matrix simplifies to:

$$\mathbf{F}_{ij} = \langle\mathcal{L}_{ij}\rangle = \frac{1}{2}\mathrm{Tr}[\mathbf{A}_i\mathbf{A}_j + \mathbf{C}^{-1}\mathbf{M}_{ij}], \tag{220}$$

where $\mathbf{A}_i \equiv \mathbf{C}^{-1}\mathbf{C}_{,i} = (\ln \mathbf{C})_{,i}$ and $\mathbf{M}_{ij} \equiv \langle\mathbf{D}_{,ij}\rangle = \boldsymbol{\mu}_{,i}\boldsymbol{\mu}^t_{,j} + \boldsymbol{\mu}_{,j}\boldsymbol{\mu}^t_{,i}$. In the limit that the data set has zero mean and is continuous rather than discrete, this equation becomes equation (101), used throughout Chapter 3.

A.3 Karhunen–Loève Eigenvalues

This section derives the Karhunen-Loève compression for data in which the mean is independent of the model parameters. This is the case in Chapter 4 where the data set has zero mean by definition. The results of this section are summarised from Tegmark, Taylor and Heavens (1997).

The Karhunen-Loève method is an example of Principal Component Analysis (and Singular Value Decomposition). Generally such methods reduce the dimensions of the data space, by rotating the data axes to point along the principal axes of the covariance matrix – i.e. the frame in which the Covariance matrix is diagonal – and then projecting onto a sub-space of these axes. The sub-space is chosen based on the corresponding eigenvalues of the axes. The highest eigenvalues correspond to the axes along which the data exhibits the greatest variance, so it is these axes which possess the greatest information.

The general form for a linear data compression is :

$$\mathbf{y} = \mathbf{B}\mathbf{x}, \tag{221}$$

where the n–dimensional data set $\mathbf{x}$ is reduced to an n'–dimensional data set $\mathbf{y}$, by the $n' \times n$ matrix $\mathbf{B}$. The compressed data set has mean and covariance:

$$\begin{aligned} \langle \mathbf{y} \rangle &= \mathbf{B}\boldsymbol{\mu} \\ \langle \mathbf{y}\mathbf{y}^t \rangle - \langle \mathbf{y} \rangle \langle \mathbf{y}^t \rangle &= \mathbf{B}\mathbf{C}\mathbf{B}^t. \end{aligned} \tag{222}$$

Substituting this into equation (220), and assuming the mean is zero, gives the Fisher matrix for the compressed data set:

$$\mathbf{F} = \frac{1}{2}\mathrm{Tr}[(\mathbf{B}\mathbf{C}\mathbf{B}^t)^{-1}(\mathbf{B}\mathbf{C}_{,i}\mathbf{B}^t)(\mathbf{B}\mathbf{C}\mathbf{B}^t)^{-1}(\mathbf{B}\mathbf{C}_{,j}\mathbf{B}^t)]. \tag{223}$$

Considering a single row of the compression matrix corresponding to compression of the data set into a single parameter, the information content can be written:

$$(2\mathbf{F}_{ii})^{1/2} = \frac{|\mathbf{b}^t\mathbf{C}_{,i}\mathbf{b}|}{\mathbf{b}^t\mathbf{C}\mathbf{b}} \tag{224}$$

where $\mathbf{b}$ is the relevant row of $\mathbf{B}$. The optimal compression is that which maximizes the information content on parameter θ_i. This is equivalent to maximizing the right hand side of

equation (224). This ratio is unchanged by multiplying $\mathbf{b}$ by a constant, so $\mathbf{b}$ can be chosen to satisfy the constraint:

$$\mathbf{b}^t\mathbf{C}\mathbf{b} = 1. \tag{225}$$

Using the method of Lagrange multipliers the optimisation problem reduces to extremizing the Lagrangean:

$$\mathbf{b}^t\mathbf{C}_{,i}\mathbf{b} - \lambda\mathbf{b}^t\mathbf{C}\mathbf{b}. \tag{226}$$

Maximising this with respect to $\mathbf{b}$ gives:

$$\mathbf{C}_{,i}\mathbf{b} = \lambda\mathbf{C}\mathbf{b} \tag{227}$$

which is a generalized eigenvalue problem. Performing a Cholesky decomposition of the covariance matrix, $\mathbf{C} = \mathbf{L}\mathbf{L}^t$ and multiplying equation (227) by $\mathbf{L}^{-1}$ from the left gives:

$$(\mathbf{L}^{-1}\mathbf{C}_{,i}\mathbf{L}^{-t})(\mathbf{L}^t\mathbf{b}) = \lambda(\mathbf{L}^t\mathbf{b}) \tag{228}$$

which is just an eigenvector–eigenvalue equation. The solution will be n orthogonal eigenvectors. The KL–compression involves sorting these eigenvalues according to their absolute value and rejecting a certain number of the smallest eigenvalue modes. Because the eigenvectors are orthogonal:

$$(\mathbf{L}^t\mathbf{b}_k)(\mathbf{L}^t\mathbf{b}_{k'}) \propto \delta_{kk'} \tag{229}$$

the compressed data satisfies:

$$\langle y_k y_{k'}\rangle = \langle(\mathbf{b}_k^t\mathbf{x})(\mathbf{x}^t\mathbf{b}_{k'})\rangle = \mathbf{b}_k^t\mathbf{C}\mathbf{b}_{k'} = \mathbf{b}_k^t\mathbf{L}\mathbf{L}^t\mathbf{b}_{k'} = \delta_{kk'} \tag{230}$$

so that the new data is statistically independent.

For a set of row vectors $\mathbf{b}$ comprising a compression matrix $\mathbf{B}$, equation (227) becomes:

$$\mathbf{C}_{,i}\mathbf{B}^t = \mathbf{C}\mathbf{B}^t\mathbf{\Lambda} \tag{231}$$

where $\mathbf{\Lambda}$ is a diagonal matrix containing the eigenvalues. Equation (230) similarly becomes:

$$\mathbf{B}\mathbf{C}\mathbf{B}^t = \mathbf{I} \tag{232}$$

so that equation (231) becomes:

$$\mathbf{B}\mathbf{C}_{,i}\mathbf{B}^t = \mathbf{\Lambda}. \tag{233}$$

From equation (223) the diagonal parts of the Fisher matrix reduce to:

$$2\mathbf{F}_{ii} = \mathrm{Tr}[\{(\mathbf{B}\mathbf{C}\mathbf{B}^t)^{-2}(\mathbf{B}\mathbf{C}_{,i}\mathbf{B}^t)\}^2] = \mathrm{Tr}\mathbf{\Lambda}^2 = \sum_{k=1}^{n'} \lambda_k^2. \tag{234}$$

The error on a parameter θ_i is then

$$\Delta\theta_i = \frac{1}{\sqrt{2}}\left(\sum_{k=1}^{n'} \lambda_k^2\right)^{-1/2} \tag{235}$$

from the Cramér–Rao inequality. This is the expression used to calculate the KL errors in Chapter 4.

B Review of the spherical harmonics methodology

This appendix reviews the formalism set out in HT. Section B.1 defines the basis functions and closure relations, Section B.2 provides derivations and definitions for the $\mathbf{\Phi}$ and $\mathbf{V}$ matrices, and Section B.3 discusses the model for small scale scattering.

B.1 Definitions of the basis functions

The function $j_l(z)$ is an l^{th}–order spherical Bessel function related to ordinary Bessel functions by $j_l(z) = \sqrt{\pi/(2z)}J_{l+\frac{1}{2}}(z)$. The spherical harmonics are defined by:

$$\begin{aligned} Y_{lm}(\theta,\phi) &= \sqrt{\frac{2l+1}{4\pi}\frac{(l-|m|)!}{(l+|m|)!}}P_l^{|m|}(\cos\theta) \\ &\quad \times \exp(im\phi) \times \begin{cases} (-1)^m & m \geq 0 \\ 1 & m < 0 \end{cases} . \end{aligned} \tag{236}$$

The boundary condition used in HT is that the radial gradient of the spherical Bessel function goes to zero at the boundary. This condition limits the possible values of k_{ln} to a discrete set. Abramowitz and Stegun (1972) give the closure relation for spherical Bessel functions whose gradient vanishes at $r = r_{max}$ as:

$$\int_0^{r_{max}} j_l(k_{lm}r)j_l(k_{ln}r)r^2dr = \frac{\pi}{4k_{ln}^3}\delta^K_{mn}\left[\frac{1}{4}+k_{ln}^2r_{max}^2-\left(l+\frac{1}{2}\right)^2\right]J^2_{l+1/2}(k_{ln}r_{max}). \tag{237}$$

It is then convenient to define the normalisation term c_{ln} as

$$c_{ln} = \frac{2k_{ln}^{3/2}}{\sqrt{\pi[1/4+k_{ln}^2r_{max}^2-(l+1/2)^2]J^2_{l+1/2}(k_{ln}r_{max})}}. \tag{238}$$

The orthogonality relation for spherical harmonics is simply:

$$\int_{4\pi} Y_{lm}(\Omega)Y^*_{l'm'}(\Omega')d\Omega = \delta^K_{ll'}\delta^K_{mm'} \tag{239}$$

and the inverse relation is:

$$\sum_{lm} Y_{lm}(\Omega)Y^*_{lm}(\Omega') = \delta_D(\hat{\mathbf{x}}-\hat{\mathbf{x}}'). \tag{240}$$

B.2 Review of the density field formalism

The observed density field is expanded into spherical Bessel functions and spherical harmonics defining our data set for the redshift survey as the set of modes:

$$\rho_\mu^s = c_\mu \int d^3s\, w(s)\rho(s) j_\mu(k_\mu s) Y_\mu^*(\theta,\phi) \tag{241}$$

which is the same as equation (169) where again s denotes the redshift space position. Continuity means that $d^3s\rho(s) = d^3r\rho(r)$. The remaining dependence on redshift is in the functions $w(s)$ and $j_\mu(k_\mu s)$. These can be expressed in terms of real space coordinates by performing a Taylor expansion about r, treating the difference $\Delta = s - r = u(r)$ as a first order perturbation:

$$w(s)j_\mu(k_\mu s) \simeq w(r)j_\mu(k_\mu r) + u(r)\frac{d}{dr}\big(w(r)j_\mu(k_\mu r)\big). \tag{242}$$

First order perturbation theory relates the radial velocity at r to the overdensity, in terms of a function of the matter density parameter $f(\Omega_m) \approx \Omega_m^{0.6}$ (Peebles 1980). The radial velocity can be written:

$$\begin{aligned} u(\mathbf{r}) &= -\hat{\mathbf{r}}\cdot\nabla\left(\frac{f(\Omega_m)}{\nabla^2}\delta_\rho(\mathbf{r})\right) \\ &= f(\Omega_m)\sum_\nu c_\nu k_\nu^{-2}\delta_\nu j_\nu'(k_\nu r)Y_\nu(\theta,\phi) \end{aligned} \tag{243}$$

where the prime indicates differentiation with respect to r. Substituting this into equations (242) and (241) gives:

$$\rho_\mu^s = (\rho_0)_\mu + \sum_\nu (b\Phi_\mu^\nu + f(\Omega_m)V_\mu^\nu)\delta_\nu, \tag{244}$$

where

$$(\rho_0)_\mu = c_\mu \int d^3r\, w(r)\rho_0(r) j_\mu(k_\mu r) Y_\mu^*(\theta,\phi) \tag{245}$$

is the transform of the weighted selection function, and the matrices $\mathbf{\Phi}$ and $\mathbf{V}$ are defined as:

$$\Phi_\mu^\nu \equiv c_\mu c_\nu \int \rho_0(\mathbf{r}) w(r) j_\mu(k_\mu r) Y_\mu^*(\theta,\phi) j_\nu(k_\nu r) Y_\nu(\theta,\phi) d^3r \tag{246}$$

and

$$V_\mu^\nu \equiv \frac{c_\mu c_\nu}{k_\nu^2}\int \rho_0(\mathbf{r})\frac{d}{dr}[w(r)j_\mu(k_\mu r)]Y_\mu^*(\theta,\phi)j_\nu'(k_\nu r)Y_\nu(\theta,\phi)d^3r. \tag{247}$$

In practice the data vector for the overdensity modes is defined as $D_\mu \equiv \rho_\mu^s - (\rho_0)_\mu$. Equation (244) is then recognised as being equivalent to equation (170). If the selection function is split into radial and angular parts:

$$\rho_0(\mathbf{r}) = M(\theta,\phi)\rho_0(r), \tag{248}$$

then the $\mathbf{\Phi}$ and $\mathbf{V}$ matrices can be similarly separated so that:

$$\Phi_\mu^\nu \equiv \Phi_\nu^\mu W_\mu^\nu \tag{249}$$

and

$$V_\mu^\nu \equiv V_\nu^\mu W_\mu^\nu \tag{250}$$

(summation not implied) with the definitions:

$$\begin{aligned}
\Phi_\mu^\nu &\equiv c_\mu c_\nu \int \rho_0(r) w(r) j_\mu(k_\mu r) j_\nu(k_\nu r) r^2 dr \\
V_\mu^\nu &\equiv \frac{c_\mu c_\nu}{k_\nu^2} \int \rho_0(r) \frac{d}{dr}[w(r) j_\mu(k_\mu r)] j_\nu'(k_\nu r) r^2 dr \\
W_\mu^\nu &\equiv \int_\Omega Y_\nu M(\Omega) Y_\mu^* d\Omega.
\end{aligned} \tag{251}$$

This separation into angular and radial parts means that each may be calculated separately allowing a faster construction of the model and convenient analysis of systematic problems. The Feldmann, Kaiser and Peacock (1994) weighting is chosen:

$$w(\mathbf{r}) = \frac{1}{1 + \rho_0(\mathbf{r}) P(k_\mu)} \tag{252}$$

where the dependence on the model power spectrum means that the weighting function depends on the mode in question.

B.3 Small scale scattering

HT also describe a correction for the small scale non–linear peculiar velocity field caused by the random motions of galaxies in virialised clusters. Their analysis involves assigning a displacement to each galaxy's radial position drawn from a probability distribution centred on the linear redshift space position. If D_μ is the overdensity in linear redshift space and D_μ' is that after taking account of small scale scatter, then the relation between the two can be written as a convolution with a scattering matrix:

$$D_\mu' = S_\mu^\nu D_\nu \tag{253}$$

where S_μ^ν is found from the following argument. If the linear redshift position s is displaced by ϵ, the final observed redshift can be written $s' = s + \epsilon$. D_μ' is then the expectation of D_μ

after averaging over all realizations, taking into account the probability distribution of ϵ. So:

$$D'_\mu = c_\mu \langle \int \rho_0(\mathbf{r}) \delta(\mathbf{r}) w(s') j_\mu(k_\mu s') Y_\mu(\Omega_{s'}) d^3 r \rangle_\epsilon. \tag{254}$$

The angular functions are unaffected by redshift distortions and so they integrate to delta functions. The final result is that the scattering matrix takes the form:

$$S_{\mu\nu} \equiv c_\mu c_\nu \Delta^K_{l_\nu l_\mu} \Delta^K_{m_\nu m_\mu} \int \int p(r-y) j_\nu(r) j_\mu(y) r dr y dy. \tag{255}$$

The probability distribution used here has an exponential form:

$$p_e(r-y) = \frac{1}{\sqrt{2}\sigma_u} \exp\left[-\frac{\sqrt{2}|r-y|}{\sigma_u}\right]. \tag{256}$$

C New parts of the spherical harmonics methodology

This appendix deals with the new formalism of the velocity field analysis. Section C.1 provides a full derivation of the velocity field noise, and Section C.2 details the practical application of the GOMA compression to the model covariance matrix.

C.1 Noise in the velocity data

The covariance of two velocity modes is

$$\begin{aligned} C_{\mu\nu} &= \left\langle \int c_\mu \left[\frac{d}{dr}[w(r)\rho(r)j_\mu(k_\mu r)] + 2\frac{w(r)\rho(r)j_\mu(k_\mu r)}{r} \right] u(r) M(\Omega) Y^*_\mu(\Omega) d^3r \right. \\ &\left. \int c_\nu \left[\frac{d}{dr'}[w(r')\rho(r')j_\nu(k_\nu r')] + 2\frac{w(r')\rho(r')j_\nu(k_\nu r')}{r'} \right] u(r') M(\Omega') Y^*_\nu(\Omega) d^3r' \right\rangle \end{aligned}$$

using expression (184). The real data set is of course a discrete sampling of both the density and velocity fields. The density function in the above equation can therefore be replaced by the selection function ρ_0 multiplied by a sum over delta functions centred at each of the galaxies' coordinates. When the integration is performed we are left with the product of two summations over the data points . The shot noise comes from those products which consider the same galaxy twice. These elements of the summation contain no information on clustering. The shot noise is therefore:

$$\begin{aligned} N_{\mu\nu} &= \left\langle \sum_i c_\mu c_\nu u^2(r_i) \mathcal{P}_\mu(Y^*_{lm}(\theta,\phi)) \mathcal{P}_\nu(Y^*_{l'm'}(\theta,\phi)) \right. \\ &\times \left[\frac{d}{dr}[w(r)\rho_0(r)j_\mu(k_\mu r)]\Big|_{r_i} + 2\frac{w(r_i)\rho(r_i)j_\mu(k_\mu r_i)}{r_i} \right] \\ &\left. \times \left[\frac{d}{dr}[w(r)\rho_0(r)j_\nu(k_\nu r)]\Big|_{r_i} + 2\frac{w(r_i)\rho(r_i)j_\nu(k_\nu r_i)}{r_i} \right] \right\rangle \end{aligned} \tag{257}$$

where as in equation (187) the data vector is split into real and imaginary parts and the functions $\mathcal{P}_{\mu,\nu}$ are correspondingly the real or imaginary parts of the spherical harmonics. A standard trick is to split the survey volume into infinitessimal cells with occupation numbers of either 0 or 1. When taking expectation values for any cell: $\langle \rho_0^2 \rangle = \langle \rho_0 \rangle$. It is also recognized that $\langle \frac{d\rho}{dr} \rangle = \frac{d\rho_0}{dr}$ and $\langle \left(\frac{d\rho}{dr}\right)^2 \rangle = \frac{d^2\rho_0}{dr^2}$. Performing the differentiation and expanding yields

equation (188). This deals with shot noise – but in a peculiar velocity survey one must also consider the scatter introduced by uncertainties in the distance indicator relation. The $D_n - \sigma$ relation at its most accurate will obtain radial velocities with a scatter of $\sigma_u^2 \approx 10\mathrm{kms}^{-1}$ but this accuracy will depend on the depth of the source in question. A simple way of incorporating this other source of noise in the above formula is to simply add this variance in quadrature to the underlying scatter. So the final expression for the noise is equation (188) with the radial velocity scatter written as the sum: $\sigma_u^2(r) = \sigma_0^2 + \sigma_{DI}^2(r)$. This incorporates both the radial velocity variance from the scatter in the Hubble relation and the velocity variance introduced by the use of inaccurate distance indicators to calculate peculiar velocities.

By following the same method as above it can be shown that there is no noise associated with the cross–power. This is because radial velocities average out at zero.

C.2 Applying GOMA to the 6dF simulations

A degenerate parameter θ_{KL} is chosen such that $\theta_{\mathrm{KL}} = ar_g + b\beta$ where a and b are parameters to be found from the Fisher matrix, which can be calculated using equation (195). For this particular analysis the values $a = 0.99$ and $b = 0.013$ are found from the eigenvectors of the Fisher matrix. The line $\theta_2 =$ constant defines the major axis of the ellipse in panel (c) of Figure 25. The derivative $\mathbf{C}_{,i}$ can be found by finite differencing over this major axis:

$$\mathbf{C}_{,i} = \frac{\mathbf{C}|_{\theta_{\mathrm{KL}}=\theta_f+\Delta\theta} - \mathbf{C}|_{\theta_{\mathrm{KL}}=\theta_f-\Delta\theta}}{2\Delta\theta} \tag{258}$$

where θ_f is the fiducial value. The full covariance matrix from Section 4.3 can be written:

$$\mathbf{C} = \frac{1}{2}\begin{pmatrix} \mathbf{\Phi P \Phi}^t + r_g\beta(\mathbf{\Phi P V}^t + \mathbf{V P \Phi}^t) + \beta^2 \mathbf{V P V}^t & r_g\beta\mathbf{\Phi P \Psi}^t + \beta^2 \mathbf{V P \Psi}^t \\ r_g\beta\mathbf{\Psi P \Phi}^t + \beta^2\mathbf{\Psi P V}^t & \beta^2\mathbf{\Psi P \Psi}^t \end{pmatrix} b^2. \tag{259}$$

Substituting this into equation (258) gives:

$$\mathbf{C}_{,i} = \frac{b^2}{2(a\Delta r_g + b\Delta\beta)} \times \begin{pmatrix} (\beta\Delta r_g + r_g\Delta\beta)(\mathbf{\Phi P V}^t + \mathbf{V P \Phi}^t) + 2\beta\Delta\beta\mathbf{V P V}^t & (\beta\Delta r_g + r_g\Delta\beta)\mathbf{\Phi P \Psi}^t + 2\beta\Delta\beta\mathbf{V P \Psi}^t \\ (\beta\Delta r_g + r_g\Delta\beta)\mathbf{\Psi P \Phi}^t + 2\beta\Delta\beta\mathbf{\Psi P V}^t & 2\beta\Delta\beta\mathbf{\Psi P \Psi}^t \end{pmatrix}$$

where $\Delta\beta = \frac{b}{a}\Delta r_g$. This can then be inserted into equation (193) to perform the Karhunen Loeve compression.